

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2004, 13:03:17 ; Search time 3909 Seconds
(without alignments)
10381.773 Million cell updates/sec

Title: US-09-990-726-222
Perfect score: 992
Sequence: 1 ggcacgagcagaactagg.....aaaaaaaaaaaaaaaaaaaaa 992

Scoring table: IDENTIFY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database :

GenEmbl :

1: gb_ba.*

2: gb_hgt.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_on.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pi.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_hgt_hum.*

31: em_hgt_inv.*

32: em_hgt_other.*

33: em_hgt_mus.*

34: em_hgt_pln.*

35: em_hgt_rod.*

36: em_hgt_nam.*

37: em_hgt_vrt.*

38: em_sv.*

39: em_hgtgo_hum.*

40: em_hgtgo_mus.*

41: em_hgtgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	992	100.0	992	6	AR252529	Sequence
2	992	100.0	992	6	AX358896	Sequence
3	992	100.0	992	6	AX362389	Sequence
4	992	100.0	992	6	AX403335	Sequence
5	504.6	50.9	505	6	BD107970	EST and e
6	501.6	50.6	502	6	AX588820	Sequence
C 7	409.8	41.3	414	11	G60502	SHGC-33533
C 8	300.4	30.3	154840	2	AC032035	Homo sapi
C 9	300.4	30.3	177738	9	AC021593	Homo sapi
C 10	300.4	30.3	209751	9	AC087645	Homo sapi
C 11	292.8	29.5	229426	2	AC010532	Homo sapi
C 12	291.2	29.4	154840	2	AC032035	Homo sapi
C 13	179.4	18.1	229426	2	AC010532	Homo sapi
C 14	163.4	16.5	116300	10	AL645856	Mouse DNA
C 15	163.4	16.5	197194	2	AL954690	Mus muscu
C 16	160.2	16.1	49532	2	AC100565	Mus muscu
C 17	149	15.0	150	11	G43545	WIAF-2405-S
C 18	149	15.0	150	11	G43546	WIAF-2405-S
C 19	127.8	12.9	161043	2	AC143361	Macaca mu
C 20	69.2	7.0	151750	9	AC003666	Homo sapi
C 21	69.2	7.0	208670	9	AC103588	Homo sapi
C 22	62.8	6.3	177921	9	AL353752	Human DNA
C 23	61.8	6.2	146883	9	AC018788	Homo sapi
C 24	61	6.1	125020	9	AF429315	Homo sapi
C 25	60.8	6.1	164762	2	AC016177	Homo sapi
C 26	60.8	6.1	170898	9	AL353764	Human DNA
C 27	60.4	6.1	129770	9	AC104684	Homo sapi
C 28	60.4	6.1	171655	2	AC067919	Homo sapi
C 29	58.8	5.9	124347	9	AC010072	Homo sapi
C 30	58.6	5.9	151993	9	AC093574	Homo sapi
C 31	58.6	5.9	194635	9	AC012087	Homo sapi
C 32	58.4	5.9	4950	9	HSM130AC2	H.sapiens m
C 33	58.4	5.9	177260	9	AC131206	Homo sapi
C 34	58	5.8	140055	2	AC023003	Homo sapi
C 35	58	5.8	151553	9	AL139342	Human DNA
C 36	58	5.8	163085	9	AC019159	Homo sapi
C 37	57.8	5.8	99268	9	AC091178	Homo sapi
C 38	57.8	5.8	119853	9	AL136374	Human DNA
C 39	57.8	5.8	138370	9	AL354713	Human DNA
C 40	57.8	5.8	156857	2	AC025375	Homo sapi
C 41	57.8	5.8	161940	2	AC079289	Homo sapi
C 42	57.4	5.8	50630	9	AC107054	Homo sapi
C 43	57.4	5.8	145380	9	AL357129	Human DNA
C 44	57.4	5.8	169202	9	AC021021	Homo sapi
C 45	57.4	5.8	175947	2	AC145171	Homo sapi
C 46	56.6	5.7	151841	9	AC013408	Homo sapi
C 47	56.6	5.7	173270	9	AC025522	Homo sapi
C 48	56.4	5.7	164405	9	AC016256	Homo sapi
C 49	56.4	5.7	166668	2	AC025258	Homo sapi
C 50	56	5.6	183778	2	AC011004	Homo sapi
C 51	56	5.6	199437	2	AC091575	Homo sapi
C 52	56	5.6	207069	9	AC090922	Homo sapi
C 53	56	5.6	254455	2	AC099902	Mus muscu
C 54	55.8	5.6	112389	2	AC073598	Homo sapi
C 55	55.8	5.6	189271	9	AL138752	Human DNA
C 56	55.4	5.6	214866	9	AC069544	Homo sapi
C 57	54.6	5.5	169986	2	AL353589	Homo sapi
C 58	54.6	5.5	181356	2	AC069441	Homo sapi
C 59	54.6	5.5	183104	9	AL355601	Human DNA
C 60	54.6	5.5	193793	9	AC099780	Homo sapi
C 61	54.6	5.5	208026	9	AF006292	Homo sapi
C 62	54.4	5.5	207190	9	AF295017	Homo sapi
C 63	54.4	5.5	217898	9	AC008498	Homo sapi
C 64	54.2	5.5	106340	9	AC097104	Homo sapi
C 65	54.2	5.5	146017	2	AC027473	Homo sapi

[illegible]

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Baker, K.P., Desnovers, L., Gerritsen, M.E., Goddard, A., Stephan, J.P.,
Godowski, P.J., Grimaldi, J.C., Gurney, A.L., Smith, V., Watanabe, C.K. and Wood, W.I.
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding
the same
JOURNAL Patent: WO 0208288-A 149 31-JAN-2002;
Genentech, Inc. (US)
FEATURES
source
1..992
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 235 a 301 c 282 g 174 t
ORIGIN
Query Match 100.0%; Score 992; DB 6; Length 992;
Best Local Similarity 100.0%; Pred. No. 1.8e-228;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCACGAGCGAGAACTAGGAGTTCTCACTGCCGAGCAGAGGCCCTACACCCACCGAG 60
Db 1 GGCACGAGCGAGAACTAGGAGTTCTCACTGCCGAGCAGAGGCCCTACACCCACCGAG 60
QY 61 GCATGGGCTCCCTGGGCTGTTCTGCTGGCGGTGCTGGCTGCCAGCGTTCTCCAAAG 120
Db 61 GCATGGGCTCCCTGGGCTGTTCTGCTGGCGGTGCTGGCTGCCAGCGTTCTCCAAAG 120
QY 121 CACGGGAGGAAGAAATACCCCTGGTCTCATTCGCTTCAAGTCTCAAGTCTGGAGTTTCC 180
Db 121 CACGGGAGGAAGAAATACCCCTGGTCTCATTCGCTTCAAGTCTGGAGTCTGGAGTTTCC 180
QY 181 CCAGAGCCGCTGGTGTCTCAATACCTGCTGTCACCCAGCCAGCCAGCCAGCCATCACT 240
Db 181 CCAGAGCCGCTGGTGTCTCAATACCTGCTGTCACCCAGCCAGCCAGCCAGCCATCACT 240
QY 241 ATTCTCTCTTGGAACCAAGAAATCAAGTGTGCGGAGAGTGTGGAAGCCAGCAGC 300
Db 241 ATTCTCTCTTGGAACCAAGAAATCAAGTGTGCGGAGAGTGTGGAAGCCAGCAGC 300
QY 301 CGGCTCTCTTCACTCAAGTCTCACTCAAGTCTCAAGTCTCAAGTCTCACTTCTCT 360
Db 301 CGGCTCTCTTCACTCAAGTCTCACTCAAGTCTCAAGTCTCAAGTCTCACTTCTCT 360
QY 361 GCGGGGCTCTCACTCAAGTCTCAAGTCTCAAGTCTCAAGTCTCAAGTCTCAAGTCT 420
Db 361 GCGGGGCTCTCACTCAAGTCTCAAGTCTCAAGTCTCAAGTCTCAAGTCTCAAGTCT 420
QY 421 AGCTGTGTCTCAAGCAGTGTCTGAGTGTGGGGCCAACTTCACTCTGAGACAGAGGG 480
Db 421 AGCTGTGTCTCAAGCAGTGTCTGAGTGTGGGGCCAACTTCACTCTGAGACAGAGGG 480
QY 481 CAGGCCCGAGGTTGAGATGATCTGCGAGGCTCTCGGAGCCAGCCAGCCATCAACACA 540
Db 481 CAGGCCCGAGGTTGAGATGATCTGCGAGGCTCTCGGAGCCAGCCAGCCATCAACACA 540
QY 541 GCCTGATCGGAAGATGGGAGTGTCCACCTGCGAGCAGAGACCTGCGAGGAGGCTG 600
Db 541 GCCTGATCGGAAGATGGGAGTGTCCACCTGCGAGCAGAGACCTGCGAGGAGGCTG 600
QY 601 CCACCTTCTCTTCTGCGAGCCAGACATCGACTGTGTTCTGTTGCGAGGCTGCAACA 660
Db 601 CCACCTTCTCTTCTGCGAGCCAGACATCGACTGTGTTCTGTTGCGAGGCTGCAACA 660
QY 661 AGCCCAATTCAGACAGCGCCCTCAAGTGGTGGCCCGAGTGGTGGCAGAGAGTGG 720
Db 661 AGCCCAATTCAGACAGCGCCCTCAAGTGGTGGCCCGAGTGGTGGCAGAGAGTGG 720
QY 721 AGGACTGGCAGGTTCCCTGGAGAGCCCATCTTGGCTTGGCTTGGCTTGGCTTGGCT 780
Db 721 AGGACTGGCAGGTTCCCTGGAGAGCCCATCTTGGCTTGGCTTGGCTTGGCTTGGCT

Db 721 AGGACTGGCAGGTTCCCTGGAGAGCCCATCTTGGCTTGGCTTGGCTTGGCTTGGCT 780
QY 781 GCGCTCTGAGTGAAGAGGAGTTTGGGGGTTTGGAGATAGGGAATGGGAGGTTCAGAGGAC 840
Db 781 GCGCTCTGAGTGAAGAGGAGTTTGGGGGTTTGGAGATAGGGAATGGGAGGTTCAGAGGAC 840
QY 841 GCAAGCAGCAGCCATGTAGTAATGAACCGTCCAGAGAGCCAGCAGCAGGAGGACTGCA 900
Db 841 GCAAGCAGCAGCCATGTAGTAATGAACCGTCCAGAGAGCCAGCAGCAGGAGGACTGCA 900
QY 901 GGCATCAGCGTCACTGTTGTAATTTGGAGTTTCATGCAAAATGAGTGTGTTTGTAGTGC 960
Db 901 GGCATCAGCGTCACTGTTGTAATTTGGAGTTTCATGCAAAATGAGTGTGTTTGTAGTGC 960
QY 961 TCTTGGCAGCAAAAAAAGAAAAAAGAAAAAAGAAAAA 992
Db 961 TCTTGGCAGCAAAAAAAGAAAAAAGAAAAAAGAAAAA 992
RESULT 4
AX403335 992 bp DNA linear PAT 14-JUN-2002
LOCUS Sequence 222 from Patent WO0073454.
DEFINITION AX403335
ACCESSION AX403335
VERSION AX403335.1 GI:21436883
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Ashkenazi, A.J., Baker, K.P., Botstein, D., Desnovers, L., Eaton, D.,
Ferrara, N., Gerber, H., Gerritsen, M., Goddard, A., Godowski, P.,
Grimaldi, C.J., Gurney, A.L., Kijavini, I., Napier, M.A., Pan, J.,
Pao, N.F., Roy, M., Stewart, T.A., Tamas, D., Watanabe, C.K.,
Williams, P., Wood, W.I. and Zhang, Z.
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding
the same
JOURNAL Patent: WO 0073454-A 222 07-DEC-2000;
Genentech Inc. (US)
FEATURES
source
1..992
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 235 a 301 c 282 g 174 t
ORIGIN
Query Match 100.0%; Score 992; DB 6; Length 992;
Best Local Similarity 100.0%; Pred. No. 1.8e-228;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCACGAGCGAGAACTAGGAGTTCTCACTGCCGAGCAGAGGCCCTACACCCACCGAG 60
Db 1 GGCACGAGCGAGAACTAGGAGTTCTCACTGCCGAGCAGAGGCCCTACACCCACCGAG 60
QY 61 GCATGGGCTCCCTGGGCTGTTCTGCTGGCGGTGCTGGCTGCCAGCGTTCTCCAAAG 120
Db 61 GCATGGGCTCCCTGGGCTGTTCTGCTGGCGGTGCTGGCTGCCAGCGTTCTCCAAAG 120
QY 121 CACGGGAGGAAGAAATACCCCTGGTCTCATTCGCTTCAAGTCTGGAGTTTCC 180
Db 121 CACGGGAGGAAGAAATACCCCTGGTCTCATTCGCTTCAAGTCTGGAGTCTGGAGTTTCC 180
QY 181 CCAGAGCCGCTGGTGTCTCAATACCTGCTGTCACCCAGCCAGCCAGCCAGCCATCACT 240
Db 181 CCAGAGCCGCTGGTGTCTCAATACCTGCTGTCACCCAGCCAGCCAGCCAGCCATCACT 240
QY 241 ATTCTCTCTTGGAACCAAGAAATCAAGTGTGCGGAGAGTGTGGAAGCCAGCAGC 300
Db 241 ATTCTCTCTTGGAACCAAGAAATCAAGTGTGCGGAGAGTGTGGAAGCCAGCAGC 300
QY 301 CGGCTCTCTTCACTCAAGTCTCACTCAAGTCTCAAGTCTCAAGTCTCACTTCTCT 360
Db 301 CGGCTCTCTTCACTCAAGTCTCACTCAAGTCTCAAGTCTCAAGTCTCACTTCTCT

301 CGGCGCTCTTCAACCTCAAGCTCACACTCAAGTCCAGTCCAGACCTGCTCACCTACTTCT 360
Db
361 GCGGGGCTCTCCACCTCAGTCCCATGTGGACAGTGCAGGCTACAGATGACTGGG 420
Qy
361 GCGGGGCTCTCCACCTCAGTCCCATGTGGACAGTGCAGGCTACAGATGACTGGG 420
Db
421 AGCTGTGTGTCACAGCCAGTGTCTGAGTGGGGGCAAACTTCACTCTGCAGGACAGAGGG 480
Qy
421 AGCTGTGTGTCACAGCCAGTGTCTGAGTGGGGGCAAACTTCACTCTGCAGGACAGAGGG 480
Db
481 CAGGCCAGGCTGGAGATGATGTCACAGGCTCTCGGGGAGCCACCTATCACCAACA 540
Qy
481 CAGGCCAGGCTGGAGATGATGTCACAGGCTCTCGGGGAGCCACCTATCACCAACA 540
Db
541 GCCTGATCGGGAAGATGGGAGGTCCACCTGCAGCAGACCATGACAGGCGAGCTG 600
Qy
541 GCCTGATCGGGAAGATGGGAGGTCCACCTGCAGCAGACCATGACAGGCGAGCTG 600
Db
601 CCAACTTCTCTCTCTCCGAGCCAGACATCGGACTGGTCTGTGTCAGGCTGCAAAACA 660
Qy
601 CCAACTTCTCTCTCTCCGAGCCAGACATCGGACTGGTCTGTGTCAGGCTGCAAAACA 660
Db
661 AGCCCAATGTCACAGCAGCCCTCAAGTGTGGGCTGCCCCAGGTGTGACCAAGATGG 720
Qy
661 AGCCCAATGTCACAGCAGCCCTCAAGTGTGGGCTGCCCCAGGTGTGACCAAGATGG 720
Db
721 AGGACTGCGAGGTCCTCTGAGAGCCCATCTTCTGCTGCGCTCTACAGGAGCACCC 780
Qy
721 AGGACTGCGAGGTCCTCTGAGAGCCCATCTTCTGCTGCGCTCTACAGGAGCACCC 780
Db
781 GCCGTCTGAGTGAAGAGAGTTTGGGGGTTTCCAGATAGGGAATGGGAGGTCAAGGAC 840
Qy
781 GCCGTCTGAGTGAAGAGAGTTTGGGGGTTTCCAGATAGGGAATGGGAGGTCAAGGAC 840
Db
841 GCAAGCAGCAGCAGTGTAGATGAACCGTCCAGAGCCCAAGCAGGACGAGTGA 900
Qy
841 GCAAGCAGCAGCAGTGTAGATGAACCGTCCAGAGCCCAAGCAGGACGAGTGA 900
Db
901 GGCCATCAGCTGACCTGCTTCTGATTTGGAGTTCATGCAAAATGAGTGTGTTAGTGC 960
Qy
901 GGCCATCAGCTGACCTGCTTCTGATTTGGAGTTCATGCAAAATGAGTGTGTTAGTGC 960
Db
961 TCTTGGCCACAAAAAATAAAAAAAAAAAAAA 992
Qy
961 TCTTGGCCACAAAAAATAAAAAAAAAAAAAA 992
Db
RESULT 5
BD107970
LOCUS
DEFINITION
EST and encoded human protein.
ACCESSION
BD107970.1 GI:23202788
VERSION
JP 2002010789-A/47.
KEYWORDS
Homo sapiens (human)
SOURCE
GENSET CORP
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 505)
Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.
EST and encoded human protein
Patent: JP 2002010789-A 47 15-JAN-2002;
TITLE
GENSET CORP
COMMENT
OS Homo sapiens (human)
PN JP 2002010789-A/47
PD 15-JAN-2002
PF 07-AUG-2000 JP 2000280989
PR 05-AUG-1999 US 60/147499
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE
GIORDANO
PC C12N15/09, C12N15/09, C07K16/47, C07K16/18, C12N1/15, C12N1/19, PC
C12N1/21,

PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC
C12N15/00
CC Von Heijne matrix
CC score 8.8999961853027
CC seq LFCLAVLAASSFS/KA
FH Key Location/Qualifiers
FT CDS 27..503
FT sig_peptide 27..80.
FEATURES
source
1..505
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 106 a 170 c 136 g 92 t 1 others
ORIGIN
Query Match 50.9%; Score 504.6; DB 6; Length 505;
Best Local Similarity 99.8%; Pred. No. 6.5e-111;
Matches 504; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 37 AGCAGAGCCCTACACCCACCGAGGCATGGGGCTCCCTGGGCTGTTCTGCTTGGCCCTGTC 96
Db 1 AGCAGAGCCCTACACCCACCGAGGCATGGGGCTCCCTGGGCTGTTCTGCTTGGCCCTGTC 60
Qy 97 TGGCTGCCACAGCTTCTCAAGGCAGGAGGAAGAAATACCCCTGCTGCTCCATTG 156
Db 61 TGGCTGCCACAGCTTCTCAAGGCAGGAGGAAGAAATACCCCTGCTGCTCCATTG 120
Qy 157 CCTACAAAGTCTTGGAAAGTTTCCCAAGGCGCTGGGTGCTCATAACTGCTGTGCAC 216
Db 121 CCTACAAAGTCTTGGAAAGTTTCCCAAGGCGCTGGGTGCTCATAACTGCTGTGCAC 180
Qy 217 CCCAGCCACACCGCCCATCACTTATTCCTCTGTGTGAAACCAAGATCAAGTGGCCA 276
Db 181 CCCAGCCACACCGCCCATCACTTATTCCTCTGTGTGAAACCAAGATCAAGTGGCCA 240
Qy 277 AGAAGTGTGTGAAGACCCAGCGCGCTCTTCAACCTCAACGTCACACTCAAGTCCA 336
Db 241 AGAAGTGTGTGAAGACCCAGCGCGCTCTTCAACCTCAACGTCACACTCAAGTCCA 300
Qy 337 GTCCAGACCTGCTCACCTACTTCTGCGGGGCTCCTCCACCTCAGGTGCCCATGTGACA 396
Db 301 GTCCAGACCTGCTCACCTACTTCTGCGGGGCTCCTCCACCTCAGGTGCCCATGTGACA 360
Qy 397 GTGCCAGGCTACAGATGCACTGAGCTGGTGTGTCAGGCTGCTGAGTGGGGCCA 456
Db 361 GTGCCAGGCTACAGATGCACTGAGCTGGTGTGTCAGGCTGCTGAGTGGGGCCA 420
Qy 457 ACTTCACCTGTCAGGACAGAGGGGCGCCCGGCTGGAGATGATCTGCCAGCGCTCCT 516
Db 421 ACTTCACCTGTCAGGACAGAGGGGCGCCCGGCTGGAGATGATCTGCCAGCGCTCCT 480
Qy 517 CGGGCAGCCACCTATCACCAACAG 541
Db 481 CGGGCAGCCACCTATCACCAACAG 505
RESULT 6
AX588820
LOCUS
DEFINITION
Sequence 695 from Patent WO02083898.
ACCESSION
AX588820
VERSION
AX588820.1 GI:27900608
KEYWORDS
Homo sapiens (human)
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
Bejanin,S., Tanaka,H., Dumas Milne Edwards,J.B., Jobert,S. and
Giordano,J.Y.
Full-length human cdnas encoding potentially secreted proteins
Patent: WO 02083898-A 695 24-OCT-2002;
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 177738)
Birren, B., Nussbaum, C. and Lander, E.
Homo sapiens chromosome 17, clone RP11-153A23
Unpublished
2 (bases 1 to 177738)
Birren, B., Linton, L., Nussbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Beda, F.,
Boguslavskiy, L., Boukhgalter, B., Brown, A., Burkett, G., Castle, A.,
Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
Dearellano, K., Dewar, K., Domino, M., Doyle, M., Fenestor, J.,
Ferrelira, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J.,
Gardyna, S., Grant, G., Hagos, B., Heatford, A., Horton, L.,
Howland, J., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Lander, T., Lehoczy, J., Levine, R., Liu, C., Liu, G., Locke, K.,
Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,
McPheters, R., Meldrim, J., Meneus, L., Morrow, J., Naylor, J.,
Norman, C.H., O'Connor, T., O'Donnell, P., Oliver, T.M., Peterson, K.,
Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,
Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W.J.,
Zimmer, A. and Zody, M.
Direct Submission
Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 177738)
Birren, B., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B.,
Camarata, J., Chang, J., Chararo, B., Choepel, Y., Collymore, A.,
Cooke, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gird, S., Graham, L., Grand-Pierre, N., Hagos, B.,
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, G., MacLean, C., Macdonald, P., Major, J., Mathews, C.,
McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mienga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,
O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (07-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 177738)
Birren, B., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B.,
Camarata, J., Chang, J., Chararo, B., Choepel, Y., Collymore, A.,
Cooke, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gird, S., Graham, L., Grand-Pierre, N., Hagos, B.,
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, G., MacLean, C., Macdonald, P., Major, J., Mathews, C.,
McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mienga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,
O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (01-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 1, 2002 this sequence version replaced gi:21702950.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L5781
Center clone name: 153_A_23
----- Location/Qualifiers
1. 177738
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-153A23"
complement(345..630)
/rpt_family="AluJo"
781..1085
/rpt_family="AluSx"
1731..1953
/rpt_family="MER20"
complement(2886..3192)
/rpt_family="AluJo"
3270..3335
/rpt_family="ORSL"
3476..3751
/rpt_family="AluSx"
complement(4214..4517)
/rpt_family="AluSp"
4554..4842
/rpt_family="AluJb"
complement(5056..5343)
/rpt_family="AluSx"
complement(5344..5641)
/rpt_family="AluSx"
complement(6335..6523)
/rpt_family="MLT1J2"
complement(6524..6821)
/rpt_family="AluSx"
complement(6822..6894)
/rpt_family="MLT1J2"
7748..7970
/rpt_family="AluSp"
8439..8735
/rpt_family="AluSx"
8736..9033
/rpt_family="AluY"
complement(9074..9373)
/rpt_family="MER46B"
9422..9487
/rpt_family="AluS"
9696..9854
/rpt_family="MER20"
9753..9761
/note="<30 qual SNGL region"
complement(10581..10677)
/rpt_family="MIR"
complement(12611..12907)
/rpt_family="AluSx"
14906..14914
/note="PCR product sequence only"
complement(15909..16192)
/rpt_family="AluSx"
17759..17801
/rpt_family="GC-rich"
17995..18021
/rpt_family="(CCG)n"
18569..18830
/rpt_family="AluSp"

```

repeat_region 18891..19237
/rpt_family="AluSx"
repeat_region 19239..19387
/rpt_family="AluY"
repeat_region 19388..19675
/rpt_family="AluSx"
repeat_region 19676..19690
/rpt_family="AluY"
repeat_region complement(20303..20438)
/rpt_family="L1MA9"
repeat_region complement(20799..21088)
/rpt_family="AluSp"
repeat_region complement(21177..21230)
/rpt_family="MIR"
repeat_region complement(22718..23024)
/rpt_family="AluJb"
repeat_region complement(23053..23147)
/rpt_family="MER53"
repeat_region complement(23349..23524)
/rpt_family="AluJo"
repeat_region 23525..23820
/rpt_family="AluSx"
repeat_region complement(23821..23936)
/rpt_family="AluJo"
repeat_region complement(23947..24088)
/rpt_family="MIR"
repeat_region 24216..24315
/rpt_family="MIR"
repeat_region complement(24316..24615)
/rpt_family="AluSx"
repeat_region 24616..24736
/rpt_family="MIR"
repeat_region complement(24742..25021)
/rpt_family="AluSx"
repeat_region 25728..26024
/rpt_family="AluSx"
repeat_region complement(26044..26353)
/rpt_family="AluSg"
repeat_region 26354..26416
/rpt_family="L1PA16"
repeat_region 26458..27006
/rpt_family="L1ME5"
repeat_region complement(27038..27252)
/rpt_family="AluSg"
repeat_region complement(27256..27450)
/rpt_family="AluSx"

Query Match 30.3%; Score 300.4; DB 9; Length 177738;
Best Local Similarity 99.7%; Pred. No. 1e-61;
Matches 301; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 131 AGAAATACCCCTGTGGTCTCCATTCCTACAAAGTCCTGGAAGTTTCCCAAGGCCG 190
Dd 25656 AGAAATACCCCTGTGGTCTCCATTCCTACAAAGTCCTGGAAGTTTCCCAAGGCCG 25597
QY 191 CTGGGTGCTATTAACCTGTGTGCACCCAGCCACCCAGCCCATCACTATTCCTCTG 250
Dd 25596 CTGGGTGCTATTAACCTGTGTGCACCCAGCCACCCAGCCCATCACTATTCCTCTG 25537
QY 251 TGGACCAAGACATCAAGTGGCCAGAGAGGTTGGTGAACCCAGCCAGCCGCTCTT 310
Dd 25536 TGGACCAAGACATCAAGTGGCCAGAGAGGTTGGTGAACCCAGCCAGCCGCTCTT 25477
QY 311 CAACCTCAAGGTCAACTCAAGTCCAGTCCAGACCTGTGCTCACTTCTGCGGCGCTC 370
Dd 25476 CAACCTCAAGGTCAACTCAAGTCCAGTCCAGACCTGTGCTCACTTCTGCGGCGCTC 25417
QY 371 CTCACCTCAGTGGCCATGTGACAGTGCCAGGCTACAGATGCACTGGAGCTGTGGTC 430
Dd 25416 CTCACCTCAGTGGCCATGTGACAGTGCCAGGCTACAGATGCACTGGAGCTGTGGTC 25357
QY 431 CA 432

```

Db 25356 CA 25355

RESULT 10
AC087645/c

LOCUS
DEFINITION
AC087645
VERSION
KEYWORDS
SOURCE

AC087645 209751 bp DNA linear PRI 05-FEB-2003
Homo sapiens chromosome 17, clone RP11-219G17, complete sequence.
AC087645
AC087645.19 GI:27311503

ORGANISM

Homo sapiens (human)
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 209751)

AUTHORS

Birren,B., Nusbaum,C. and Lander,E.

TITLE

Homo sapiens chromosome 17, clone RP11-219G17

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 209751)

AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B., Brown,A.,
Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
Collumore,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.,
Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T.,
Lehoczky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P.,
Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
McPheters,R., Meldrum,J., Meneus,L., Mihova,T., Mienga,V.,
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Poillara,V., Raymond,C., Retta,R.,
Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Sougniez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Travers,M., Travis,N., Triglio,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.

TITLE

Direct Submission

JOURNAL

Submitted (15-JAN-2001) Whitehead Institute/MIT Center for Genome

REFERENCE

Research, 320 Charles Street, Cambridge, MA 02141, USA

AUTHORS

3 (bases 1 to 209751)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collumore,A.,
Cooke,P., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J.,
Matthews,C., McCarthy,M., Meldrum,J., Meneus,L., Mihova,T.,
Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., North,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Tophann,K.,
Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE

Direct Submission

JOURNAL

Submitted (16-JAN-2003) Whitehead Institute/MIT Center for Genome

REFERENCE

Research, 320 Charles Street, Cambridge, MA 02141, USA

AUTHORS

4 (bases 1 to 209751)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collumore,A., Cooke,P., Cooke,P., Corum,B., DeArelano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,


```
Db 181052 TGAACCAAGACATCAAGTGGCCAAAGAGGTGGTGAAGACCAAGCGCGCTCCTT 180993
QY 311 CAACCTCAACGTCACACTCAAGTCCAGTCCAGACCTGCTCCTCACTACTTCTGCGGGCGTC 370
Db 180992 CAACCTCAACGTCACACTCAAGTCCAGTCCAGACCTGCTCCTCACTACTTCTGCGGGCGTC 180933
QY 371 CTCACCTCAGTGGCCCATGTCGACATGTCGACAGTCCAGGCTACAGATGACATGCGAGCTGTGCTC 430
Db 180932 CTCACCTCAGTGGCCCATGTCGACATGTCGACAGTCCAGGCTACAGATGACATGCGAGCTGTGCTC 180873
QY 431 CA 432
Db 180872 CA 180871

RESULT 11
AC010532 229426 bp DNA linear HTG 05-SEP-2000
LOCUS Homo sapiens chromosome 17 clone RP11-219G17, LOW-PASS SEQUENCE
DEFINITION SAMPLING.
ACCESSION AC010532
VERSION AC010532.2 GI:6758797
KEYWORDS HTG; HTGS PHASE0.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 229426)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 17
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 229426)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT On Jan 26, 2000 this sequence version replaced gi:5882406.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----
* NOTE: This record contains 99 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
1 563: contig of 563 bp in length
* 1 564 1436: contig of 873 bp in length
* 1437 1649: contig of 213 bp in length
* 1650 2523: contig of 874 bp in length
* 2524 3286: contig of 763 bp in length
* 3287 3613: contig of 327 bp in length
* 3614 5430: contig of 1817 bp in length
* 5431 6681: contig of 1251 bp in length
* 6682 7357: contig of 676 bp in length
* 7358 8117: contig of 760 bp in length
* 8118 8718: contig of 601 bp in length
* 8719 9464: contig of 746 bp in length
* 9465 9964: contig of 500 bp in length
* 9965 10959: contig of 995 bp in length
* 10960 11873: contig of 914 bp in length
* 11874 12656: contig of 783 bp in length
* 12657 13674: contig of 1018 bp in length
* 13675 14935: contig of 1261 bp in length
* 14936 16423: contig of 1488 bp in length
* 16424 17186: contig of 763 bp in length
* 17187 18079: contig of 893 bp in length
* 18080 18813: contig of 734 bp in length
* 18814 19729: contig of 916 bp in length
* 19730 20675: contig of 946 bp in length
* 20676 21721: contig of 1046 bp in length
* 21722 23111: contig of 1390 bp in length
* 23112 25188: contig of 2077 bp in length
* 25189 26588: contig of 1400 bp in length
* 26589 27644: contig of 1056 bp in length
* 27645 29372: contig of 1728 bp in length
* 29373 29495: contig of 123 bp in length
* 29496 30431: contig of 936 bp in length
* 30432 31449: contig of 1018 bp in length
* 31450 32983: contig of 1534 bp in length
* 32984 33162: contig of 179 bp in length
* 33163 33975: contig of 813 bp in length
* 33976 35031: contig of 1056 bp in length
* 35032 35709: contig of 678 bp in length
* 35710 37308: contig of 1599 bp in length
* 37309 38107: contig of 799 bp in length
* 38108 39102: contig of 995 bp in length
* 39103 39591: contig of 489 bp in length
* 39592 41151: contig of 1560 bp in length
* 41152 42842: contig of 1691 bp in length
* 42843 43828: contig of 986 bp in length
* 43829 44009: contig of 181 bp in length
* 44010 45000: contig of 999 bp in length
```

```

* 44010 44874: contig of 865 bp in length
*      gap of unknown length
* 44875 45191: contig of 317 bp in length
*      gap of unknown length
* 45192 46147: contig of 956 bp in length
*      gap of unknown length
* 46148 47504: contig of 1357 bp in length
*      gap of unknown length
* 47505 49091: contig of 1587 bp in length
*      gap of unknown length
* 49092 50259: contig of 1168 bp in length
*      gap of unknown length
* 50260 51361: contig of 1102 bp in length
*      gap of unknown length
* 51362 52631: contig of 1270 bp in length
*      gap of unknown length
* 52632 53607: contig of 976 bp in length
*      gap of unknown length
* 53608 54820: contig of 1213 bp in length
*      gap of unknown length
* 54821 56600: contig of 1780 bp in length
*      gap of unknown length
* 56601 58236: contig of 1636 bp in length
*      gap of unknown length
* 58237 59716: contig of 1480 bp in length
*      gap of unknown length
* 59717 62221: contig of 2505 bp in length
*      gap of unknown length
* 62222 65085: contig of 2864 bp in length
*      gap of unknown length
* 65086 66739: contig of 1654 bp in length
*      gap of unknown length
* 66740 68359: contig of 1620 bp in length
*      gap of unknown length
* 68360 70014: contig of 1655 bp in length
*      gap of unknown length
* 70015 71808: contig of 1794 bp in length
*      gap of unknown length
* 71809 73581: contig of 1773 bp in length
*      gap of unknown length
* 73582 75574: contig of 1993 bp in length
*      gap of unknown length
* 75575 78468: contig of 2894 bp in length
*      gap of unknown length
* 78469 79231: contig of 763 bp in length
*      gap of unknown length
* 79232 83240: contig of 4009 bp in length
*      gap of unknown length
* 83241 86157: contig of 2917 bp in length
*      gap of unknown length
* 86158 89150: contig of 2993 bp in length
*      gap of unknown length
* 89151 92671: contig of 3521 bp in length
*      gap of unknown length
* 92672 95442: contig of 2771 bp in length
*      gap of unknown length
* 95443 98476: contig of 3034 bp in length
*      gap of unknown length
* 98477 100815: contig of 2339 bp in length
*      gap of unknown length
* 100816 103005: contig of 2190 bp in length
*      gap of unknown length
* 103006 105514: contig of 2509 bp in length
*      gap of unknown length
* 105515 108704: contig of 3190 bp in length
*      gap of unknown length
* 108705 111097: contig of 2393 bp in length
*      gap of unknown length
* 111098 114888: contig of 3791 bp in length
*      gap of unknown length
* 114889 117423: contig of 2535 bp in length
*      gap of unknown length

Query Match      29.5%; Score 292.8; DB 2; Length 229426;
Best Local Similarity 97.7%; Pred. No. 6.8e-60;
Matches 297; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 688 CAGTGTGCCCCCAGGTGGTGACGACGAGAGATGGAGGACTGCAGGAGTCCCTCGAGAGCC 747
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 748 CCATCCTTTGCTTCCCGCTCTACAGGAGCACCCCGCTCTCAGTCAAGAGGAGTTTGGGG 807
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 135356 CCATCCTTTGCTTCCCGCTCTACAGGAGCACCCCGCTCTCAGTCAAGAGGAGTTTGGGG 135415
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 808 GGTTCAGGATAGGGAATGGGGAGGTGAGAGACGCAAGCAAGCAGCAGCCATGTAGATGAAC 867
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 135416 GGTTCAGGATAGGGAATGGGGAGGTGAGAGACGCAAGCAAGCAGCAGCCATGTAGATGAAC 135475
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 868 CGTCCAGAGAGCCCAAGCAGGCGAGGACTGCAGGCCATCAGCGTGCACCTGTCGTATTT 927
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 928 GGAGTTTCATGCAAAATGAGTGTGTTTAGTGCTCTTGGCCACAAAAAAGGAGGAGGAGG 987
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 135536 GGAGTTTCATGCAAAATGAGTGTGTTTAGTGCTCTTGGCCACAAAAAAGGAGGAGGAGG 135595
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 988 AAAA 991
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
AC032035
LOCUS      154840 bp      DNA      linear      HTG 08-SEP-2000
DEFINITION Homo sapiens chromosome 17 clone RP11-141D15 map 17, *** SEQUENCING
IN PROGRESS ***, 18 unordered pieces.
AC032035
ACCESSION      AC032035.3 GI:9994161
VERSION
KEYWORDS      HTG; HTGS PHASE1.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 154840)
AUTHORS      Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE      Homo sapiens chromosome 17, clone RP11-141D15
JOURNAL
REFERENCE
  2 (bases 1 to 154840)
AUTHORS      Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
Boguslavskiy, L., Boukhgalter, B., Brown, A., Burkett, G.,
Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
Collins, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Gage, D.,
Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Galagan, J., Gardyna, S., Glinde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J.,
Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,
Meldrim, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neil, D., Ollivar, T.M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Triggilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (03-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 8, 2000 this sequence version replaced 91:7705196.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)

```


http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L9138
 Center clone name: 141_D_15

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 18 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

1 2373: contig of 2373 bp in length
* 2374 2473: gap of 100 bp
* 2474 4930: contig of 2457 bp in length
* 4931 5030: gap of 100 bp
* 5031 7215: contig of 2185 bp in length
* 7216 7315: gap of 100 bp
* 7316 9541: contig of 2226 bp in length
* 9542 9641: gap of 100 bp
* 9642 12211: contig of 2570 bp in length
* 12212 12311: gap of 100 bp
* 12312 14923: contig of 2612 bp in length
* 14924 15023: gap of 100 bp
* 15024 19094: contig of 4071 bp in length
* 19095 19194: gap of 100 bp
* 19195 22286: contig of 3092 bp in length
* 22287 22386: gap of 100 bp
* 22387 26147: contig of 3761 bp in length
* 26148 26247: gap of 100 bp
* 26248 31933: contig of 5686 bp in length
* 31934 32033: gap of 100 bp
* 32034 37339: contig of 5306 bp in length
* 37340 37439: gap of 100 bp
* 37440 42858: contig of 5419 bp in length
* 42859 42958: gap of 100 bp
* 42959 48810: contig of 5852 bp in length
* 48811 48910: gap of 100 bp
* 48911 56029: contig of 7119 bp in length
* 56030 56129: gap of 100 bp
* 56130 62189: contig of 6060 bp in length
* 62190 82959: contig of 20670 bp in length
* 82960 83059: gap of 100 bp
* 83060 107059: contig of 24000 bp in length
* 107060 107159: gap of 100 bp
* 107160 154840: contig of 47681 bp in length.
  
```

FEATURES

source

```

1. 154840
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /chromosome="17"
  /map="17"
  /clone="RP11-141D15"
  /clone_lib="RP11-11 Human Male BAC"
  BASE COUNT 36180 a 41741 c 41557 g 32874 t 2488 others
  ORIGIN
  
```

```

Query Match 29.4%; Score 291.2; DB 2; Length 154840;
Best Local Similarity 97.4%; Pred. No. 1.7e-59;
Matches 296; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
  
```

```

QY 688 CAGTGTGCCCCCAGGTGTGTACCAAGATGGAGACTGGAGGTCCTCCCTGGAGAGCC 747
  |||||
  
```

```

Db 86135 CACTGCTTTGTCAAGTGTGTGACCAAGATGGAGACTGGAGGTCCTCCCTGGAGAGCC 86194
  |||||
  
```

```

QY 748 CCATCTTCCTTGGCGCTCTACAGAGACACCGCGCTCTGAGTGAAGAGAGTTTGGGG 807
  |||||
Db 86195 CCATCTTTCCTTGGCGCTCTACAGAGACACCGCGCTTGTAGTGAAGAGAGTTTGGGG 86254
  |||||
QY 808 GCTTCAGATAGGGAATGGGAGGTTCAGAGGACGCAAGCAGCAGCATGTAGATGAAC 867
  |||||
Db 86255 GCTTCAGATAGGGAATGGGAGGTTCAGAGGACGCAAGCAGCAGCATGTAGATGAAC 86314
  |||||
QY 868 COTCCAGAGGCCAAGCAGCGGAGGACTGCGAGGCATCAGGCGCATGTCGTATTT 927
  |||||
Db 86315 COTCCAGAGGCCAAGCAGCGGAGGACTGCGAGGCATCAGGCGCATGTCGTATTT 86374
  |||||
QY 928 GGAGTTTCATGCAAAATGAGTGTGTTTACCTCTTGCACAAAAAATAAAAAA 987
  |||||
Db 86375 GGAGTTTCATGCAAAATGAGTGTGTTTACCTCTTGCACAAAAAATAAAAAA 86434
  |||||
QY 988 AAAA 991
  |||||
Db 86435 AAAA 86438
  |||||

RESULT 13
AC010532/c
DEFINITION Homo sapiens chromosome 17 clone RP11-219G17, LOW-PASS SEQUENCE
LOCUS AC010532
ACCESSION AC010532
VERSION AC010532.2 GI:6758797
KEYWORDS HTG; HTGS; PHASE0.
SOURCE Homo sapiens
ORGANISM Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 229426)
  DOE Joint Genome Institute.
  Sequencing of Human Chromosome 17
  Unpublished
  2 (bases 1 to 229426)
  DOE Joint Genome Institute.
  Direct SubtilBion
  Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint
  Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
  On Jan 26, 2000 this sequence version replaced gi:5882406.
  -----Genome Center
  Center: Joint Genome Institute
  Center Code: JGI
  Web site: http://www.jgi.doe.gov
  -----
  * NOTE: This record contains 99 individual
  * sequencing reads that have not been assembled into
  * contigs. Runs of N are used to separate the reads
  * and the order in which they appear is completely
  * arbitrary. Low-pass sequence sampling is useful for
  * identifying clones that may be gene-rich and allows
  * overlap relationships among clones to be deduced.
  * However, it should not be assumed that this clone
  * will be sequenced to completion. In the event that
  * the record is updated, the accession number will
  * be preserved.
  * 1 563: contig of 563 bp in length
  * gap of unknown length
  * 564 1436: contig of 873 bp in length
  * gap of unknown length
  * 1437 1649: contig of 213 bp in length
  * gap of unknown length
  * 1650 2523: contig of 874 bp in length
  * gap of unknown length
  * 2524 3286: contig of 763 bp in length
  * gap of unknown length
  * 3287 3613: contig of 327 bp in length
  * gap of unknown length
  * 3614 5430: contig of 1817 bp in length
  * gap of unknown length
  
```

* 5431	6681:	contig of 1251 bp in length	* 42843	43828:	gap of unknown length
* 6682	7357:	gap of unknown length	* 43829	44009:	contig of 986 bp in length
* 7358	8117:	gap of unknown length	* 44010	44874:	gap of unknown length
* 8118	8718:	contig of 601 bp in length	* 44875	45191:	contig of 865 bp in length
* 8719	9464:	gap of unknown length	* 45192	46147:	gap of unknown length
* 9465	9964:	gap of unknown length	* 46148	47504:	contig of 956 bp in length
* 9965	10959:	gap of unknown length	* 47505	49091:	contig of 1357 bp in length
* 10960	11873:	gap of unknown length	* 49092	50259:	contig of 1587 bp in length
* 11874	12656:	gap of unknown length	* 50260	51361:	gap of unknown length
* 12657	13674:	gap of unknown length	* 51362	52631:	contig of 1102 bp in length
* 13675	14935:	gap of unknown length	* 52632	53607:	gap of unknown length
* 14936	16423:	gap of unknown length	* 53608	54820:	contig of 976 bp in length
* 16424	17186:	contig of 1018 bp in length	* 54821	56600:	gap of unknown length
* 17187	18079:	gap of unknown length	* 56601	58236:	contig of 1213 bp in length
* 18080	18813:	contig of 893 bp in length	* 58237	59716:	gap of unknown length
* 18814	19729:	gap of unknown length	* 59717	62221:	contig of 1780 bp in length
* 19730	20675:	gap of unknown length	* 62222	65085:	contig of 1636 bp in length
* 20676	21721:	gap of unknown length	* 65086	66739:	gap of unknown length
* 21722	23111:	contig of 1046 bp in length	* 66740	68359:	contig of 1480 bp in length
* 23112	25188:	gap of unknown length	* 68360	70014:	gap of unknown length
* 25189	26588:	contig of 1390 bp in length	* 70015	71808:	contig of 1655 bp in length
* 26589	27644:	gap of unknown length	* 71809	73581:	gap of unknown length
* 27645	29372:	contig of 1056 bp in length	* 73582	75574:	contig of 1794 bp in length
* 29373	29495:	gap of unknown length	* 75575	78468:	gap of unknown length
* 29496	30431:	contig of 123 bp in length	* 78469	79231:	gap of unknown length
* 30432	31449:	gap of unknown length	* 79232	83240:	contig of 763 bp in length
* 31450	32983:	contig of 1018 bp in length	* 83241	86157:	gap of unknown length
* 32984	33162:	gap of unknown length	* 86158	89150:	contig of 4009 bp in length
* 33163	33975:	contig of 1534 bp in length	* 89151	92671:	contig of 2917 bp in length
* 33976	35031:	gap of unknown length	* 92672	95442:	gap of unknown length
* 35032	35709:	gap of unknown length	* 95443	98476:	gap of unknown length
* 35710	37308:	contig of 678 bp in length	* 98477	100815:	contig of 3034 bp in length
* 37309	38107:	gap of unknown length	* 100816	103005:	gap of unknown length
* 38108	39102:	contig of 1599 bp in length	* 103006	105514:	contig of 2339 bp in length
* 39103	39591:	gap of unknown length	* 105515	108704:	gap of unknown length
* 39592	41151:	contig of 799 bp in length	* 108705	111097:	contig of 2771 bp in length
* 41152	42842:	gap of unknown length			gap of unknown length
		contig of 1560 bp in length			
		gap of unknown length			
		contig of 1691 bp in length			

```

* 111098 114888: contig of 3791 bp in length
*      *      gap of unknown length
* 114889 117423: contig of 2535 bp in length
*      *      gap of unknown length

Query Match
Best Local Similarity 18.1%; Score 179.4; DB 2; Length 229426;
Matches 183; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 244 CCTCTGTGAACCAAGATCAATCAAGTGGCCAAAGAGTGGTGAAGACCCACGAGCCGG 303
Db 1331 CGTTCGGGGAACCAAGATCAATCAAGTGGCCAAAGAGTGGTGAAGACCCACGAGCCGG 1272
QY 304 CTTCTTTCAACCTCAAGTCACTCAAGTCCAGTCCAGACCTGCTCACTACTTTGACC 363
Db 1271 CTTCTTTCAACCTCAAGTCACTCAAGTCCAGTCCAGACCTGCTCACTACTTTGACC 1212
QY 364 GGGGCTCTCCACCTCAAGTGGCCATGTGACAGTGGCAGGCTACAGATGCATGGGAGC 423
Db 1211 GGGGCTCTCCACCTCAAGTGGCCATGTGACAGTGGCAGGCTACAGATGCATGGGAGC 1152
QY 424 TGTGTGTCCA 432
Db 1151 TGTGTGTCCA 1143

RESULT 14
AL645856
LOCUS
DEFINITION
  Mouse DNA sequence from clone RP23-151N19 on chromosome 11,
  complete sequence.
ACCESSION
  AL645856
VERSION
  AL645856.5 GI:22331400
KEYWORDS
  HTG.
SOURCE
  Mus musculus (house mouse)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 116300)

REFERENCE
  1 (bases 1 to 116300)
AUTHORS
  White, S.
TITLE
  Direct Submission
JOURNAL
  Submitted (24-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,
  Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
  humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
  On Aug 27, 2002 this sequence version replaced gi:21738435.

COMMENT
  ----- Genome Center
  Center: Wellcome Trust Sanger Institute
  Center code: SC
  Web site: http://www.sanger.ac.uk
  Contact: humquery@sanger.ac.uk

  -----
  During sequence assembly data is compared from overlapping clones.
  Where differences are found these are annotated as variations
  together with a note of the overlapping clone name. Note that the
  variation annotation may not be found in the sequence submission
  corresponding to the overlapping clone, as we submit sequences with
  only a small overlap as described above.
  This sequence was finished as follows unless otherwise noted: all
  regions were either double-stranded or sequenced with an alternate
  chemistry or covered by high quality data (i.e., phred quality >=
  30); an attempt was made to resolve all sequencing problems, such
  as compressions and repeats; all regions were covered by at least
  one plasmid subclone or more than one M13 subclone; and the
  assembly was confirmed by restriction digest. The following
  abbreviations are used to associate primary accession numbers given
  in the feature table with their source databases: Em: EMBL; Sw:
  SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP
  database can be found at
  http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-151N19 is
  from the RBC1-23 Mouse PAC Library
  constructed by the group of Pieter
  de Jong.
  For further details see http://www.chori.org/bacpac/home.htm
  VECTOR: pBac3.6.

```

```

FEATURES
  source
  Location/Qualifiers
  1..116300
  /organism="Mus musculus"
  /mol_type="genomic DNA"
  /db_xref="taxon:10090"
  /chromosome="11"
  /clone="RP23-151N19"
  /clone_lib="RPCI-23"
  BASE COUNT 29288 28679 29788 28545 2
  ORIGIN
  Query Match 16.5%; Score 163.4; DB 10; Length 116300;
  Best Local Similarity 73.3%; Pred. No. 1.1e-28;
  Matches 209; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

  QY 148 TCTCCATTGCTCAACAAGTCTCTGGAAGTTTCCCAAGAGCGCGCTGGTGTCTATAACCT 207
  Db 93425 TCACCATGCTCAACAAGTACTGGAAGTTTATCCCAAGCGGAGGGTCTTATAACCT 93484
  QY 208 GCTGTGCACCCACAGCCACCCATCACCTATTTCCTCTCTGTGGAACCAAGAACATCA 267
  Db 93485 GCGATGCCCTGAGCGTCCCAGCCCATCACATCTCTCTCTGGTAGCGGAGTATCC 93544
  QY 268 AGGTGCCAAGAAGTGGTGAAGACCCACGAGCGCGCTCTCTTCAACCTCAACGTACAC 327
  Db 93545 TGGTGGCAAAAAGGTTGTGCATGACTCCGTGCGCGCTCTCTTCAACATCATATCA 93604
  QY 328 TCAAGTCCAGTCCAGACCTGCTCACCTACTTCTGCGGGGCTCTCCACTCAGGTGCC 387
  Db 93605 TCAAGTCCAGTCCAGACCTGCTCACCTACTTCTGCGGGGCTCTCCACTCAGGTGCC 93664
  QY 388 ATGTGGACAGTGCAGGCTACAGATGCATCTGAGGCTGTGGTCCA 432
  Db 93665 ATGGACCCAGCAGCGCTCCAGATGTACAGGAAGTGTGGGCTA 93709

  RESULT 15
  AL954690
  LOCUS
  DEFINITION
    Mus musculus chromosome 11 clone RP23-386A4, *** SEQUENCING IN
    PROGRESS ***.
  ACCESSION
    AL954690
  VERSION
    AL954690.2 GI:25136861
  KEYWORDS
    HTG; HTGS PHASE2; HTGS CANCELLED.
  SOURCE
    Mus musculus (house mouse)
  ORGANISM
    Mus musculus
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
    1 (bases 1 to 197194)

  REFERENCE
    1 (bases 1 to 197194)
  AUTHORS
    Tromans, A.
  TITLE
    Direct Submission
  JOURNAL
    Submitted (13-NOV-2002) Wellcome Trust Sanger Institute, Hinxton,
    Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
    humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
    On Nov 19, 2002 this sequence version replaced gi:25045706.
    Sequence from the Mouse Genome Sequencing Consortium whole genome
    shotgun may have been used to confirm this sequence. Sequence data
    from the whole genome shotgun alone has only been used where it has
    a phred quality of at least 30.

  COMMENT
    ----- Genome Center
    Center: Wellcome Trust Sanger Institute
    Center code: SC
    Web site: http://www.sanger.ac.uk
    Contact: humquery@sanger.ac.uk
    ----- Project Information
    Center project name: BM386A4
    ----- Summary Statistics
    Assembly program: XGAP4; version 4.5
    Chemistry: Dye-terminator; 100% of reads
    Consensus quality: 195990 bases at least Q40
    Consensus quality: 196234 bases at least Q30
    Consensus quality: 196338 bases at least Q20
    Insert size: 197194; sum-of-contigs

```

```

Insert size: 189681; 3.6% error; agarose-fp
Quality coverage: 7.24x in Q20 bases; sum-of-contigs Quality
coverage: 7.53x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
    source
        1..197194
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
            /chromosome="11"
            /clone="RP23-386A4"
            /clone_lib="RPCI-23"
            /note="assembly_fragment:02398"
            /clone_end:SP6
            vector_side:left
    misc_feature
        1..197194
            /note="assembly_fragment:02398"
            /clone_end:SP6
            vector_side:left
    BASE COUNT
        46441 a 51277 c 51550 g 47922 t 4 others
    ORIGIN
        Query Match      16.5%; Score 163.4; DB 2; Length 197194;
        Best Local Similarity 73.3%; Pred. No. 1.1e-28;
        Matches 209; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 148 TCTCATTTGCTTACAAAGTCTCGGAAGTTTCCCAAGGCGGTGGTGTCTCAAACT 207
Db 20120 TCACCATTTGCTTACAAAGTCTCGGAAGTTTCCCAAGGCGGTGGTGTCTCAAACT 20179

QY 208 GCTGTGACCCAGCCAGCCAGCCATCTATTCCTCTGTGGACCAAGACATCA 267
Db 20180 GCGATGCCCCCTGAGCGCTCCAGCCCATCACATCTCTCTCTGCTAGCCGAGTATCC 20239

QY 268 AGGTGGCCAAAGTGTGTGAAGACCCAGAGCCGGCTCTTCAACCTCAACGTCACAC 327
Db 20240 TGTGGCAAAAAGTGTGTGATGACTCGTGGCCGGCTCTTCAACATCATATCA 20299

QY 328 TCAAGTCCAGTCCAGACTGTCTACCTACTTGTCCGGCGGTCTCCACCTCAGTGCC 387
Db 20300 TCAAGTCCAGCCAGACCTGTCTACCTACTCTGCGAGGCACTCGAAGCTTGGCACT 20359

QY 388 ATGTGGACAGTCCAGGCTACAGATGACACTGGAGCTGTGTCCA 432
Db 20360 ATGGACCCAGCAGCGAGGTCCAGATGTACCGAAGTGTGGCTA 20404

RESULT 16
AC100565 49532 bp DNA linear HTG 22-NOV-2001
LOCUS Mus musculus clone RP23-154K21, LOW-PASS SEQUENCE SAMPLING.
DEFINITION AC100565
ACCESSION AC100565
VERSION AC100565.1 GI:17047931
KEYWORDS HTG; HTGS PHASE0.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 49532)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Mus musculus, clone RP23-154K21
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 49532)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguelavkiy,L., Boukhalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepe,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Illiev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,

```

```

Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Truberman,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,U., Zimmer,A. and Zody,W.
Direct Submission
Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L15700
Center clone name: 154_K21
-----
* NOTE: This record contains 48 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
1 986: contig of 986 bp in length
* 987 1086: gap of 100 bp
* 1087 2009: contig of 923 bp in length
* 2010 2109: gap of 100 bp
* 2110 3007: contig of 898 bp in length
* 3008 3107: gap of 100 bp
* 3108 3866: contig of 759 bp in length
* 3867 3966: gap of 100 bp
* 3967 4874: contig of 908 bp in length
* 4875 4974: gap of 100 bp
* 4975 5839: contig of 865 bp in length
* 5840 5939: gap of 100 bp
* 5940 7017: contig of 1078 bp in length
* 7018 7117: gap of 100 bp
* 7118 8079: contig of 962 bp in length
* 8080 8179: gap of 100 bp
* 8180 9139: contig of 960 bp in length
* 9140 9239: gap of 100 bp
* 9240 10254: contig of 1015 bp in length
* 10255 10354: gap of 100 bp
* 10355 11289: contig of 935 bp in length
* 11290 11389: gap of 100 bp
* 11390 12334: contig of 945 bp in length
* 12335 12434: gap of 100 bp
* 12435 13410: contig of 976 bp in length
* 13411 13510: gap of 100 bp
* 13511 14445: contig of 935 bp in length
* 14446 14545: gap of 100 bp
* 14546 15272: contig of 727 bp in length
* 15273 15373: gap of 100 bp
* 15373 16394: contig of 1022 bp in length
* 16395 16494: gap of 100 bp
* 16495 17461: contig of 967 bp in length
* 17462 17561: gap of 100 bp

```

```

* 17562 18545: contig of 984 bp in length
* 18546 18645: gap of 100 bp
* 18646 19600: contig of 955 bp in length
* 19601 19700: gap of 100 bp
* 20670 20770: contig of 970 bp in length
* 20771 21739: contig of 969 bp in length
* 21740 21839: gap of 100 bp
* 21840 22669: contig of 830 bp in length
* 22670 22769: gap of 100 bp
* 22770 23763: contig of 994 bp in length
* 23764 24820: contig of 957 bp in length
* 24821 24920: gap of 100 bp
* 24921 25847: contig of 927 bp in length
* 25848 26787: contig of 840 bp in length
* 26788 26887: gap of 100 bp
* 26888 27831: contig of 944 bp in length
* 27832 27931: gap of 100 bp
* 27932 28881: contig of 950 bp in length
* 28882 28981: gap of 100 bp
* 28982 29906: contig of 925 bp in length
* 29907 30921: contig of 915 bp in length
* 30922 31021: gap of 100 bp
* 31022 31757: contig of 736 bp in length
* 31758 31857: gap of 100 bp
* 31858 32798: contig of 941 bp in length
* 32799 32898: gap of 100 bp
* 32899 33835: contig of 937 bp in length
* 33836 34860: contig of 925 bp in length
* 34861 34960: gap of 100 bp
* 34961 35887: contig of 927 bp in length
* 35888 35987: gap of 100 bp
* 35988 36990: contig of 1003 bp in length
* 36991 37090: gap of 100 bp
* 37091 38085: contig of 995 bp in length
* 38086 38185: gap of 100 bp
* 38186 39172: contig of 987 bp in length
* 39173 39272: gap of 100 bp
* 39273 40215: contig of 943 bp in length
* 40216 40315: gap of 100 bp
* 40316 41267: contig of 952 bp in length
* 41268 41367: gap of 100 bp
* 41368 42271: contig of 904 bp in length
* 42272 42371: gap of 100 bp
* 42372 43302: contig of 931 bp in length
* 43303 43402: gap of 100 bp
* 43403 44283: contig of 881 bp in length
* 44284 44383: gap of 100 bp
* 44384 45395: contig of 1012 bp in length
* 45396 45495: gap of 100 bp
* 45496 46390: contig of 895 bp in length
* 46391 46490: gap of 100 bp
* 46491 47408: contig of 918 bp in length
* 47409 47508: gap of 100 bp
* 47509 48438: contig of 930 bp in length
* 48439 48538: gap of 100 bp
* 48539 49532: contig of 994 bp in length.
FEATURES
  source
    Location/Qualifiers
      1..49532
        /organism="Mus musculus"
        /mol_type="genomic DNA"
        /db_xref="taxon:10090"
        /clone="RP23-154K21"
        /clone_lib="RPCI-23 Female Mouse BAC"
BASE COUNT 11202 a 11441 c 10738 g 11027 t 5124 others
ORIGIN
  Query Match 16.1%; Score 160.2; DB 2; Length 49532;
  Best Local Similarity 72.6%; Pred. No. 6.6e-28;

```

```

Matches 207; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
QY 148 TCCTCATTTGCTACAAAGTCTCTGGAAGTTTCCCAAGCCGCTGGTGTCTATAACCT 207
      |||
Db 9291 TCACCTTTGCTACAAAGTACTGGAAGTTTATCCCAAGCCGAGGGTCTTATAACCT 9350
      |||
QY 208 GCTGTGACCCCGCCAGCCACCGCCATCACTTCCCTCTGTGGAACCAAGACATCA 267
      |||
Db 9351 GCGATGCCCTTGAGCGCTCCAGCCCATCATCTCTCTCTCTGCTAGCGGAGGTATCC 9410
      |||
QY 268 AGGTGGCCAAAGAGTGGTGAACACCCAGCGCGCTCTCTTCAACCTCAAGCTCACAC 327
      |||
Db 9411 TGGTGGCAAAAAGAGTTTGCATGACTCCGTGGCGGCTCTTCAACATCAATACCA 9470
      |||
QY 328 TCAAGTCCAGTCAGACCTGCTCCTACTTTCGCGGGGCTCTCACTCACTAGGTGCC 387
      |||
Db 9471 TCAAGTCCAGCCAGACCTGCTCCTACTCTCTGCGAGGCACTCGAACTCTGGCACCT 9530
      |||
QY 388 ATGTGCACAGTGCAGGCTACAGATGCATGGGAGCTGTGGTCCA 432
      |||
Db 9531 ATGGACCCAGCAGCGCTCCAGATGTACAGAACTGTGGGCTA 9575

RESULT 17
LOCUS G43545/c
DEFINITION WIAP-2405-STS Human Thudson EST Homo sapiens STS cDNA, sequence
          tagged site.
ACCESSION G43545
VERSION G43545.1 GI:4192462
KEYWORDS STS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 150)
AUTHORS Wang,D.G., Fan,J.B., Siao,C.J., Berno,A., Young,P., Sapolsky,R.,
          Ghandour,G., Perkins,N., Winchester,E., Spencer,J., Kruglyak,L.,
          Stein,L., Hsie,L., Topaloglou,T., Hubbell,E., Robinson,E.,
          Mittmann,M., Morris,M.S., Shen,N., Kilburn,D., Rioux,J.,
          Nusbaum,C., Rozen,S., Hudson,T.J., Lipshutz,R., Chee,M. and
          Lander,E.S.
TITLE Large-scale identification, mapping, and genotyping of
          single-nucleotide polymorphisms in the human genome
JOURNAL Science 280 (5366), 1077-1082 (1998)
MEDLINE 98248615
PUBMED 9582121
COMMENT Synonyms: EST376246a, EST376246
          Contact: Thomas Hudson
          Whitehead Institute/MIT Center for Genome Research
          Whitehead Institute for Biomedical Research
          9 Cambridge Center, Cambridge MA 02142 USA
          Tel: 617 252 1900
          Fax: 617 252 1902
          Email: thudson@genome.wi.mit.edu
          Primer A: GTGGCAAGACGAGCTAAAC
          Primer B: GGAATGGGAGGTCAGAGG
          STS size: 150
          PCR Profile:
            Presoak: 94 degrees C for 4.00 minutes
            Denaturation: 94 degrees C for 50.0 seconds
            Annealing: 58 degrees C for 1.50 minutes
            Polymerization: 72 degrees C for 1.00 minutes
            PCR Cycles: 30
            Thermal Cycler: custom built by IAS, Costar, Cambridge MA

Protocol:
  Template: 10 ng
  Primer: each 5 pM
  dNTPs: 4 mM
  Taq Polymerase: 0.5 U
  Total Vol: 20 uL
Buffer:

```

```

Mg2+: 1.5 mM
KCl: 50 mM
Tris-HCl: 10 mM
Gelatin: .001 %
Location/Qualifiers
1. .150
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="Human Thudson EST"
/note="STS derived from sequences in dbEST and the
Unigene collection."
1. .150
1. .20
primer_bind 30 a 46 c 33 g 40 t 1 others
BASE COUNT 30 a 46 c 33 g 40 t 1 others
ORIGIN
Query Match 15.0%; Score 149; DB 11; Length 150;
Best Local Similarity 99.3%; Pred. No. 3.8e-25;
Matches 149; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 820 GGAATGGGAGGTCAGAGGACGCAAGCAGCAGCATGTAAGTGAACCGTCCAGAGAGC 879
Db 150 GGAATGGGAGGTCAGAGGACGCAAGCAGCAGCATGTAAGTGAACCGTCCAGAGAGC 91
QY 880 CAAGCAGCGCAGAGACTGCAGGCGATCAGCGTGCACCTGTCGATTGAGTTCATGCA 939
Db 90 CAAGCAGCGCAGAGACTGCAGGCGATCAGCGTGCACCTGTCGATTGAGTTCATGCA 31
QY 940 AAATGAGTGTGTTTAGTCTCTTGGCCAC 969
Db 30 AAATGAGTGTGTTTAGTCTCTTGGCCAC 1

RESULT 18
G43546/c
LOCUS
DEFINITION
WAF-2407-STS Human Thudson EST Homo sapiens STS cdna, sequence
tagged site.
G43546
ACCESSION
VERSION G43546.1 GI:4192463
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 150)
Wang, D.G., Fan, J.B., Siao, C.J., Berno, A., Young, P., Sapolsky, R.,
Ghandour, G., Perkins, N., Winchester, E., Spencer, J., Kruglyak, L.,
Stein, L., Hsie, L., Topaloglou, T., Hubbell, E., Robinson, E.,
Mittmann, M., Morris, M.S., Shen, N., Kilburn, D., Rioux, J.,
Nusbaum, C., Rozen, S., Hudson, T.J., Lipnütz, R., Chee, M. and
Lander, E.S.
Large-scale identification, mapping, and genotyping of
single-nucleotide polymorphisms in the human genome
Science 280 (5366), 1077-1082 (1998)
98248615
9582121
Synonyms: EST376246b, EST376246
Contact: Thomas Hudson
Whitehead Institute/MIT Center for Genome Research
Whitehead Institute for Biomedical Research
9 Cambridge Center, Cambridge MA 02142 USA
Tel: 617 252 1900
Fax: 617 252 1902
Email: thudson@genome.wi.mit.edu
Primer A: GTGGCAAGCAGCTAAAC
Primer B: GGAATGGGAGGTCAGAGG
STS size: 150
PCR Profile:
Presoak: 94 degrees C for 4.00 minutes
Denaturation: 94 degrees C for 50.0 seconds

```

```

Annealing: 58 degrees C for 1.50 minutes
Polymerization: 72 degrees C for 1.00 minutes
PCR Cycles: 30
Thermal Cycler: custom built by IAS, Costar, Cambridge MA

Protocol:
Template: 10 ng
Primer: each 5 pM
dNTPs: 4 mM
Taq Polymerase: 0.5 U
Total Vol: 20 uL

```

```

Buffer:
Mg2+: 1.5 mM
KCl: 50 mM
Tris-HCl: 10 mM
Gelatin: .001 %
Location/Qualifiers
1. .150
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="Human Thudson EST"
/note="STS derived from sequences in dbEST and the
Unigene collection."
1. .150
1. .20
primer_bind 30 a 47 c 32 g 40 t 1 others
BASE COUNT 30 a 47 c 32 g 40 t 1 others
ORIGIN

```

```

FEATURES
source
1. .150
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="Human Thudson EST"
/note="STS derived from sequences in dbEST and the
Unigene collection."
1. .150
1. .20
primer_bind 30 a 47 c 32 g 40 t 1 others
BASE COUNT 30 a 47 c 32 g 40 t 1 others
ORIGIN
Query Match 15.0%; Score 149; DB 11; Length 150;
Best Local Similarity 99.3%; Pred. No. 3.8e-25;
Matches 149; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 820 GGAATGGGAGGTCAGAGGACGCAAGCAGCAGCATGTAAGTGAACCGTCCAGAGAGC 879
Db 150 GGAATGGGAGGTCAGAGGACGCAAGCAGCAGCATGTAAGTGAACCGTCCAGAGAGC 91
QY 880 CAAGCAGCGCAGAGACTGCAGGCGATCAGCGTGCACCTGTCGATTGAGTTCATGCA 939
Db 90 CAAGCAGCGCAGAGACTGCAGGCGATCAGCGTGCACCTGTCGATTGAGTTCATGCA 31
QY 940 AAATGAGTGTGTTTAGTCTCTTGGCCAC 969
Db 30 AAATGAGTGTGTTTAGTCTCTTGGCCAC 1

```

```

RESULT 19
AC143361
LOCUS
DEFINITION
Macaca mulatta clone CH250-271N12, *** SEQUENCING IN PROGRESS ***.
ACCESSION
AC143361
VERSION
AC143361.1 GI:29648548
KEYWORDS
HTG; HTGS PHASE2; HTGS PGI.
SOURCE
Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
Cercopitheciinae; Macaca.
1 (bases 1 to 161043)
Csuros, M. and Milosavljevic, A.
Pooled genomic indexing (PGI): mathematical analysis and experiment
design
(in) Guigo, R. and Gusfield, D. (Eds.);
ALGORITHMS IN BIOINFORMATICS, SECOND INTERNATIONAL WORKSHOP, WABI
2002, ROME, ITALY, SEPTEMBER 17-21, 2002, PROCEEDINGS: 10-28;
Springer (2002)
2 (bases 1 to 161043)
Milosavljevic, A., Sodergren, E., Csuros, M., Li, B., Jackson, A.R.,
Adams, C., Adio-Oduola, B., Ali-ouman, F.R., Allen, C., Alsbrooks, S.L.,
Anaratunge, H.C., Are, J.R., Ayele, M., Banks, T., Barbara, J.,
Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D., Bouck, J.,
Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C.,

```

```

* the accession number will be preserved.
*
FEATURES             Source
1..161043
   Location/Qualifiers
       /organism="Macaca mulatta"
       /mol_type="genomic DNA"
       /db_xref="taxon:9544"
       /clone="CH250-271N12"
1..161043
   note="assembly.name:CH250-271N12.1B
   CONFIDENCE: 0.83"
BASE COUNT      2845 a   3627 c   3686 g   2810 t 148075 others
ORIGIN
Query Match              12.9%;   Score 127.8;   DB 2;   Length 161043;
Best Local Similarity    76.3%;   Pred. No. 4.2e-20;
Matches 135;   Conservative 0;   Mismatches 42;   Indels 0;   Gaps 0;

QY      529  CTATCACCACACGCGCTGATCGGGAAGGATGGCGAGGTCCACCTGCAGCAGACGACCATGCC 588
Db      1   CGANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNAGGTCCATCTCCAGCAGCAGCACGCG 60

QY      589  ACAGCGAGCGTCCCAACTTCTCTCTCTCCGAGCGACGACATCGGACTGGTCTTGTTGTCGCC 648
Db      61  ACGGGCAGCGTCCCAACTTCTCTCTCTCCGCGGGCAGACATCGGACTGGTTCAGTGGC 120

QY      649  AGGCTGCAAAACAACGCCAATGTCACGACAGCGCCCTCACAGTGGTGCCCGCCAGGTG 705
Db      121  AGGCTGCAAAACAAGCGTCAATGTCCAGCAGCGCCCTCACATGGTGCCCGCCAGGTG 177

RESULT 20
LOCUS      AC003666
DEFINITION Homo sapiens Xp22 BAC GS-551019 (Genome Systems Human BAC library)
            and cosmids U199A7 and U209F2 (Lawrence Livermore X chromosome
            cosmid library) containing part of human chloride channel 4 gene,
            complete sequence.
ACCESSION  AC003666
VERSION     AC002358
KEYWORDS    AC002360 AC003017
SOURCE      HTG.
            AC003666.1  GI:2992476
ORGANISM    Homo sapiens (human)
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 151750)
            Muzny, D., Aronson, A.D., Brundage, B., Carvelli, K., Chen, E., Chen, J.,
            Di.W., Ding, Y., Dugan, S., Durbin, J., Forcum, J., Ganesi, R.,
            Garcia, C., Goodman, M., Gorrell, J.H., Haywood, M., Jackson, L.,
            Jin, S., Kampal, R., Karpathy, S., Leal, B., Li, Y., Liu, W., Logan, O.,
            Lu, J., Ly, T., Martinez, C., Oswal, G., Perez, L., Rashid, N.D.,
            Rowland, K., Savage, L., Scherer, S.S., Shen, H., Timms, K.M., Todd, J.,
            Vo.Q., Worley, K.C., Yu, W., Chinault, C., Nelson, D. and Gibbs, R.A.
            Direct Submission
            Unpublished
            2 (bases 1 to 151750)
            Worley, K.C.
            Direct Submission
            Submitted (09-DEC-1997) Molecular and Human Genetics, Baylor
            College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
            3 (bases 1 to 151750)
            Worley, K.C.
            Direct Submission
            Submitted (27-MAR-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            4 (bases 1 to 151750)
            Worley, K.C.
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 151750)
            Direct Submission
            Submitted (05-MAY-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77
```

AUTHORS TITLE JOURNAL

Worley, K.C.
Direct Submission
Submitted (30-JUL-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
6 (bases 1 to 151750)

REFERENCE AUTHORS TITLE JOURNAL

Worley, K.C.
Direct Submission
Submitted (07-AUG-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
Sequencing is completed to a minimum standard of double strand
coverage with a minimum of 2 clones and 2 reads with no ambiguities
or 2 chemistries with a minimum of 2 clones and 3 reads with no
ambiguities. If the sequence quality does not meet this standard,
it will be indicated in the annotation.

COMMENT

The repeat regions shown were identified using RepeatMasker by
Adrian Smit.
Sequence similarities were identified using PowerBlast by Jinghui
Zhang.

Exon/Intron boundaries of identified genes were chosen if there
were canonical splice junctions that maintained sequence continuity
across the splice junctions.

FEATURES source

1. 151750
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xp22"
/clone="GS-551019, U199A7, U209F2"
/clone_lib="Genome Systems Human BAC library, Lawrence
Livermore X chromosome library"

misc_feature

1. 374
/note="Overlap bases 16683 to 167256 in AC008008"
/function="Overlap with clone AC008008"

repeat_region

complement(2657..2689)

repeat_region

/rpt_family="AT rich"

repeat_region

complement(2773..2863)

repeat_region

/rpt_family="MERSB"

repeat_region

3472..4871

repeat_region

/rpt_family="L1PA2"

repeat_region

5843..6143

repeat_region

/rpt_family="L1MA3"

repeat_region

complement(6989..7493)

repeat_region

/rpt_family="L1MA3"

repeat_region

complement(7513..7811)

repeat_region

/rpt_family="AluSx"

STS

11128..11201

repeat_region

/rpt_family="GGAA)n"

STS

11487..11618

repeat_region

/standard_name="hz-1"

repeat_region

11757..11981

repeat_region

/rpt_family="MER20"

repeat_region

12660..12782

repeat_region

/rpt_family="L2"

repeat_region

12879..13195

repeat_region

/rpt_family="MLTIE"

repeat_region

complement(15364..15659)

repeat_region

/rpt_family="AluJc"

repeat_region

complement(15936..16099)

repeat_region

/rpt_family="MER20"

repeat_region

complement(16717..17085)

repeat_region

/rpt_family="L1MB7"

repeat_region

complement(17159..17374)

repeat_region

/rpt_family="L1MB7"

repeat_region

complement(17421..19544)

repeat_region

/rpt_family="L1M4"

repeat_region

complement(19654..19760)

repeat_region

/rpt_family="L1MB7"

repeat_region

complement(20189..20669)

repeat_region

/rpt_family="L1MB7"

repeat_region

20670..21171

repeat_region

/rpt_family="MLTIF"

repeat_region

complement(21172..21792)

repeat_region

/rpt_family="L1MB7"

repeat_region

22005..22430

repeat_region

/rpt_family="L1MCI"

repeat_region

23332..23445

repeat_region

/rpt_family="L1PA4"

repeat_region

23469..23588

repeat_region

/rpt_family="FLAM A"

repeat_region

complement(25042..25067)

repeat_region

/rpt_family="POLY A"

repeat_region

complement(25163..25265)

repeat_region

/rpt_family="L1PA5"

repeat_region

25264..25431

repeat_region

/rpt_family="L1PA5"

repeat_region

complement(25569..26063)

repeat_region

/rpt_family="MER31A"

repeat_region

complement(27469..27570)

repeat_region

/rpt_family="MIR"

repeat_region

28608..28643

repeat_region

/rpt_family="TAAAA)n"

repeat_region

complement(28713..29010)

repeat_region

/rpt_family="AluJb"

repeat_region

29209..29310

repeat_region

/rpt_family="MIR"

repeat_region

29312..29613

repeat_region

/rpt_family="AluSx"

repeat_region

30127..30683

repeat_region

/rpt_family="MLTID"

repeat_region

31072..31160

repeat_region

/rpt_family="(GAA)n"

repeat_region

31161..31279

repeat_region

/rpt_family="(GGA)n"

repeat_region

complement(31772..31891)

repeat_region

/rpt_family="(GA)n"

repeat_region

32352..32572

repeat_region

/rpt_family="MIR"

repeat_region

complement(33455..33479)

repeat_region

/rpt_family="AT rich"

repeat_region

34912..35382

repeat_region

/rpt_family="L1MA1"

repeat_region

35405..35701

repeat_region

/rpt_family="AluSg"

repeat_region

35707..36176

repeat_region

/rpt_family="L1MA1"

repeat_region

36178..38428

STS

36898..37069

STS

/standard_name="STS1-CSRL-31b6-uA/CSRL-31b6-uZ, Chr. 7,

STS

/db_xref="dbSTS:7821"

STS

37958..38089

STS

/standard_name="A002D07, Chr. 7, Homo sapiens"

STS

/db_xref="dbSTS:32826"

STS

38436..38566

STS

/rpt_family="L1MA1"

STS

38597..38795

STS

/rpt_family="L1PB3"

STS

complement(38769..40304)

STS

/rpt_family="L1PA12"

STS

40309..40609

STS

/rpt_family="AluY"

STS

complement(40626..40905)

STS

/rpt_family="L1PA12"

STS

40904..41028

STS

/rpt_family="L1PB3"

STS

complement(41977..42037)

STS

/rpt_family="AT rich"

STS

complement(42391..42718)

STS

/rpt_family="MER34"


```

repeat_region 42979..43067
/rpt_family="L2"
repeat_region complement(43085..43220)
/rpt_family="MIR"
repeat_region 43445..43746
/rpt_family="AluSx"
repeat_region 44344..44380
/rpt_family="MIR"
repeat_region complement(45352..45584)
/rpt_family="MIR"
repeat_region complement(46643..46707)
/rpt_family="L2"
repeat_region complement(46905..47293)
/rpt_family="MSTA"

Query Match
Best Local Similarity 7.0%; Score 69.2; DB 9; Length 151750;
Matches 131; Conservative 0; Mismatches 58; Indels 7; Gaps 2;

QY 796 AGGAGTTGGGGGTTTCAGGATAGGATGGGAGGTGAGAGGCGCAAAAGCAGCAGCA 855
Db |||||
QY 5846 AGGAGTTAACAGGGGTGGATAGGAGTGGAGTGGAGTACGAGTACGAGGACGCGTA 5905
Db |||||
QY 856 TGTAGATGAACCGTCACAGAGAGCAAGCAGGCA---GAGGACTGCAGGCCATCAGCGT 912
Db |||||
QY 5906 TGTAGATGAACAAATCTGGAGATCTAACACACACACCTGAGGACTACATGGCATAAATTT 5965
Db |||||
QY 913 GCACGTGTCGTATTTGGAGTTCATGCAAAAGCAGTGTCTTTAGTGTCTCTGCACAAA 972
Db |||||
QY 5966 ATACTGA---ATTGGGATTCATAAGAAATGAGTAGATTTTAGTGTCTCTGCACAAA 6021
Db |||||
QY 973 AAAAAAAAAAAAAAAAAA 988
Db |||||
QY 6022 AATAAAGCAAAAAA 6037
Db |||||

RESULT 21
AC103588 208670 bp DNA linear PRI 21-MAR-2002
LOCUS AC103588 Homo sapiens chromosome 3 clone RP11-640H5, complete sequence.
DEFINITION AC103588
ACCESSION AC103588
VERSION AC103588.2 GI:19570161
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 208670)
AUTHORS Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 208670)
AUTHORS Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
Saenphimmachak,C., Phelps,K.A., Raymond,C. and Haugen,E.D.
Direct Submission
Unpublished
2 (bases 1 to 208670)
AUTHORS Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
Saenphimmachak,C., Phelps,K.A., Raymond,C. and Haugen,E.D.
Direct Submission
Submitted (29-NOV-2001) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
3 (bases 1 to 208670)
AUTHORS Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
Saenphimmachak,C., Phelps,K.A., Raymond,C. and Haugen,E.D.
Direct Submission
Submitted (21-MAR-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
On Mar 21, 2002 this sequence version replaced gi:17149452.
COMMENT ----- Genome Center
Center: University of Washington Genome Center
Center Code: UWGC
Web site: http://www.genome.washington.edu
Contact: uwgchgs@u.washington.edu
----- Project Information
Center project name: chr-3
Center clone name: RP11-640H5 (bc0507)
----- Summary Statistics

```

Sequencing vector: plasmid; 108752; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 208594 bases at least Q40
 Consensus quality: 208666 bases at least Q30
 Consensus quality: 208670 bases at least Q20
 Insert size: 207271; sum-of-contigs
 Quality coverage: 7.7x in Q20 bases; sum-of-contigs

Overlapping Sequences:

5': Mapping in progress
 3': RP11-1C5 (UWGC:bc0101) AC093411, 107464-bp overlap

Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp.
 Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted:
 all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

EcoRI				BglII				HindIII			
SeqDerMap	EngPrnt	SeqDerMap	EngPrnt	SeqDerMap	EngPrnt	SeqDerMap	EngPrnt	SeqDerMap	EngPrnt	SeqDerMap	EngPrnt
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
8696	8812	2524	2527	561	<800	561	<800	561	<800	561	<800
6	<800	2067	2075	6382	6445	6382	6445	6382	6445	6382	6445
1516	1544	9401	9533	512	<800	512	<800	512	<800	512	<800
5561	5606	17986	18248	449	<800	449	<800	449	<800	449	<800
10971	10860	4692	4675	2484	2483	2484	2483	2484	2483	2484	2483
2516	2558	2347	2350	11918	12069	11918	12069	11918	12069	11918	12069
1834	1834	7017	7123	511	<800	511	<800	511	<800	511	<800
2909	2978	3961	3953	1038	1087	1038	1087	1038	1087	1038	1087
1008	1008	1756	1735	4818	4777	4818	4777	4818	4777	4818	4777
2704	2761	6058	5990	1677	1673	1677	1673	1677	1673	1677	1673
13906	13758	5740	5669	2657	2610	2657	2610	2657	2610	2657	2610
1875	1834	4526	4503	1348	1312	1348	1312	1348	1312	1348	1312
3723	3725	3178	3199	462	<800	462	<800	462	<800	462	<800

abbreviations are used to associate primary accession numbers given in the feature table with their source databases: En: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr9> RP11-497K15 is from the library RP11-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBACE3.6 This sequence is the entire insert of clone RP11-497K15 The true left end of clone RP11-40C6 is at 128937 in this sequence. The true right end of clone RP11-5802 is at 80971 in this sequence.

FEATURES

```

1. .177921
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="9"
/clone="RP11-497K15"
/clone_lib="RPC1-11.2"
7. .372
/note="match: GSS: Em:AQ521699"
7. .187
/note="match: GSS: Em:AQ400784"
507. .633
/note="L12 repeat: matches 2370. .2494 of consensus"
649. .751
/note="L1MB8 repeat: matches 6074. .6173 of consensus"
1133. .1428
/note="AluX repeat: matches 1. .297 of consensus"
3216. .3644
/note="MSTB repeat: matches 1. .420 of consensus"
6009. .6477
/note="MLTID repeat: matches 3. .501 of consensus"
6702. .6982
/note="Aluub repeat: matches 1. .308 of consensus"
7119. .7203
/note="MIR repeat: matches 81. .167 of consensus"
7337. .8076
/note="match: GSS: Em:AQ752801"
9330. .8613
/note="L12 repeat: matches 2465. .2750 of consensus"
8633. .8927
/note="L12 repeat: matches 2088. .2408 of consensus"
9107. .9266
/note="L12 repeat: matches 2593. .2750 of consensus"
9700. .10001
/note="AluSc repeat: matches 1. .303 of consensus"
10683. .10783
/note="MIR repeat: matches 31. .145 of consensus"
10890. .10969
/note="MIR repeat: matches 81. .159 of consensus"
complement(11756. .12346)
/note="match: STS: Em:HSJ20A4"
11931. .11992
/note="L12 repeat: matches 2688. .2750 of consensus"
12476. .12860
/note="MLTID repeat: matches 106. .514 of consensus"
13210. .13318
/note="FLAM A repeat: matches 1. .109 of consensus"
complement(13812. .14276)
/note="match: GSS: Em:B58243"
13898. .14064
/note="FRAM repeat: matches 0. .168 of consensus"
14484. .14783
/note="AluX repeat: matches 1. .301 of consensus"
16037. .16512
/note="match: GSS: Em:AQ515148"
16045. .16482
/note="AluX repeat: matches 284. .284 of consensus"
/note="L12 repeat: matches 302. .628 of consensus"
20983. .21292
/note="AluY repeat: matches 1. .309 of consensus"
21321. .21428
/note="FLAM A repeat: matches 21. .128 of consensus"
22740. .22830
/note="MIR repeat: matches 30. .125 of consensus"
22890. .22940
/note="L12 repeat: matches 2700. .2749 of consensus"
complement(23361. .23469)
/note="match: STS: Em:U67307"
24070. .24362
/note="AluSq repeat: matches 6. .296 of consensus"
24403. .25216
/note="L1MA6 repeat: matches 5487. .6295 of consensus"
25217. .25355
/note="L1PB1 repeat: matches 6011. .6151 of consensus"
25356. .25420
/note="L1MA6 repeat: matches 5425. .5487 of consensus"
25415. .25744
/note="L1MA6 repeat: matches 5091. .5403 of consensus"
26967. .27168
/note="MIR repeat: matches 49. .262 of consensus"
27190. .27327
/note="Charliel repeat: matches 2620. .2761 of consensus"
27323. .27475
/note="Charliel repeat: matches 64. .226 of consensus"
27607. .27917
/note="AluX repeat: matches 1. .306 of consensus"
28407. .29843
/note="L1MB7 repeat: matches 4591. .6173 of consensus"
32140. .32448
/note="AluX repeat: matches 1. .304 of consensus"
33717. .34022
/note="AluSc repeat: matches 1. .306 of consensus"
34170. .34419
/note="MLTIG repeat: matches 31. .260 of consensus"
34527. .34626
/note="MLTIF repeat: matches 440. .541 of consensus"
35085. .35128
/note="BUR1 repeat: matches 4967. .5010 of consensus"
35525. .35836
/note="AluX repeat: matches 1. .312 of consensus"
35844. .35911
/note="L1PB2 repeat: matches 349. .409 of consensus"
36507. .36760
/note="L1PB2 repeat: matches 4128. .4385 of consensus"
36761. .37068
/note="AluY repeat: matches 1. .308 of consensus"
37069. .38794
/note="L1PB2 repeat: matches 4385. .6105 of consensus"
38795. .39098
/note="Aluub repeat: matches 2. .302 of consensus"
39099. .39146
/note="L1PB2 repeat: matches 6105. .6152 of consensus"
39612. .39895
/note="AluX repeat: matches 284. .284 of consensus"

```

```

repeat_region 40251..40597
/note="L2 repeat: matches 2390..2747 of consensus"
repeat_region 40814..41296
/note="MLR1D repeat: matches 1..505 of consensus"
repeat_region 43064..43491
/note="MLR2FB repeat: matches 1..412 of consensus"
repeat_region 43800..43934
/note="MLR1H repeat: matches 412..543 of consensus"
repeat_region 44369..44772
/note="MSTA repeat: matches 3..426 of consensus"
repeat_region 45202..45287
/note="MIR repeat: matches 115..208 of consensus"
repeat_region 45352..45855
/note="MER1A repeat: matches 1..523 of consensus"
repeat_region 46156..46517
/note="THB1B repeat: matches 1..363 of consensus"
repeat_region 47513..47824
/note="AluSx repeat: matches 1..312 of consensus"
repeat_region 48695..48996
/note="AluSc repeat: matches 1..304 of consensus"
repeat_region 49341..49553
/note="MIR repeat: matches 6..252 of consensus"
misc_feature complement(49807..50346)
/note="match: GSS: Em:AQ174373"

Query Match 6.3%; Score 62.8; DB 9; Length 177921;
Best Local Similarity 61.6%; Pred. No. 0.0002;
Matches 117; Conservative 0; Mismatches 72; Indels 1; Gaps 1;

QY 802 TTGGGGGTTTCAGGATAGGGAATGGGAGTTCAGAGCAGCGACTGCGCCATCAGCTGCTTC 861
|||||
Db 76569 TTGAGGGAAGACCAATGGGAGATGTAGTTCAGAGGACACAAAGTAGCAGACGTGTAGG 76628

QY 862 ATGAACTTCAGAGAGCCAGCAGCGAGGACTGCGCCATCAGCTGCTTC 921
|||||
Db 76629 ATGAACAAACTAGAGACTAATGTACACATGATTTCTAGAGTAATAAATTGTACT- 76687

QY 922 GTATTGGAGTTTCAGAAATAGTGTGTTTGTAGCTGCTTCGCCACAAAAA 981
|||||
Db 76688 GTATTGGAGTTTCGCAATTAGTAGATTCTAGTGCCTTCGCCACAGAAAGGAAA 76747

QY 982 AAAAAA 991
|||||
Db 76748 AAATGATAA 76757

RESULT 23
AC018788/c
LOCUS AC018788 146883 bp DNA linear PRI 22-NOV-2001
DEFINITION Homo sapiens chromosome , clone RP11-24E9, complete sequence.
ACCESSION AC018788
VERSION AC018788.10 GI:16117587
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 146883)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Beckerly,R., Bada,F.,
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cooke,P.,
Dearellano,K., Dewar,K., Domino,M., Doyle,N., Fenesfor,J.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Grand-Pierre,N., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
Lamarez,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrum,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Oct 13, 2001 this sequence version replaced gi:16041478.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information
Center project name: L4443
Center clone name: 24_E_9

-----
FEATURES
source 1..146883
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="RP11-24E9"
/clone_lib="RP11-24E9"
/complement(455..723)
/rpt_family="MIR"
2337..2409
/rpt_family="L2"
/rpt_family="L2"
complement(2418..2815)
/rpt_family="L3"
complement(3143..3415)
/rpt_family="L2"
complement(3428..3588)
/rpt_family="L3"
4076..4116
/rpt_family="MIR"
complement(4701..4992)
/rpt_family="AluSg"
5007..5205
/rpt_family="MIR"
5300..5371
/rpt_family="L2"

```

```

O'Connor,T., O'Donnell,P., Olivari,T.M., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D., Roy,A.,
Santos,R., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J., Zimmer,A.
and Zody,M.
Direct Submission
Submitted (19-DEC-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 146883)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cooke,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
Lamarez,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrum,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Oct 13, 2001 this sequence version replaced gi:16041478.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information
Center project name: L4443
Center clone name: 24_E_9

-----
FEATURES
source 1..146883
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="RP11-24E9"
/clone_lib="RP11-24E9"
/complement(455..723)
/rpt_family="MIR"
2337..2409
/rpt_family="L2"
/rpt_family="L2"
complement(2418..2815)
/rpt_family="L3"
complement(3143..3415)
/rpt_family="L2"
complement(3428..3588)
/rpt_family="L3"
4076..4116
/rpt_family="MIR"
complement(4701..4992)
/rpt_family="AluSg"
5007..5205
/rpt_family="MIR"
5300..5371
/rpt_family="L2"

```

```

repeat_region complement(7086..7283)
repeat_region /rpt_family="MER5A"
repeat_region 7418..7464
repeat_region /rpt_family="L2"
repeat_region complement(8426..8712)
repeat_region /rpt_family="LIPAl0"
repeat_region complement(8753..8917)
repeat_region /rpt_family="MER54A"
repeat_region complement(8984..9076)
repeat_region /rpt_family="MIR"
repeat_region 9639..9748
repeat_region /rpt_family="MIR"
repeat_region complement(9766..10614)
repeat_region /rpt_family="MER54A"
repeat_region 10700..10834
repeat_region /rpt_family="L2"
repeat_region 10859..10967
repeat_region /rpt_family="GA-rich"
repeat_region complement(11035..11307)
repeat_region /rpt_family="L3"
repeat_region 11316..11349
repeat_region /rpt_family="(TC)n"
repeat_region 11349..11391
repeat_region /rpt_family="(CA)n"
repeat_region complement(11869..11975)
repeat_region /rpt_family="MLT1G1"
repeat_region complement(12252..14205)
repeat_region /rpt_family="LIPAS"
repeat_region complement(14965..15444)
repeat_region /rpt_family="MER114"
repeat_region 15495..15722
repeat_region /rpt_family="LIME1"
repeat_region 15819..16227
repeat_region /rpt_family="LIP3"
repeat_region 16228..16256
repeat_region /rpt_family="(TTTG)n"
repeat_region 16257..16678
repeat_region /rpt_family="LIP3"
repeat_region 16679..17746
repeat_region /rpt_family="MER11C"
repeat_region 17747..19574
repeat_region /rpt_family="LIP3"
repeat_region 19575..21107
repeat_region /rpt_family="LIP3"
repeat_region complement(24551..24629)
repeat_region /rpt_family="MIR"
repeat_region complement(24706..25065)
repeat_region /rpt_family="LIPB1"
repeat_region 25066..25181
repeat_region /rpt_family="(TA)n"
repeat_region 25183..25866
repeat_region /rpt_family="LI"
repeat_region 25871..25976
repeat_region /rpt_family="(TA)n"
repeat_region complement(25991..26108)
repeat_region /rpt_family="LIP5"
repeat_region complement(26175..26288)
repeat_region /rpt_family="LIPAl5-16"
repeat_region complement(26305..26969)
repeat_region /rpt_family="LIPAl5-16"
repeat_region 26964..27157
repeat_region /rpt_family="LI"
repeat_region complement(27164..27230)
repeat_region /rpt_family="MER61D"
repeat_region 27296..27400
repeat_region /rpt_family="LIPAl6"
repeat_region complement(27401..27697)
repeat_region /rpt_family="AluSp"
repeat_region 27698..29368
repeat_region /rpt_family="LIPAl6"
unsure 29271..29293
repeat_region /note="<30 qual SNGL region"
repeat_region complement(29369..29564)

```

```

/rpt_family="MER61D"
29981..30691
/rpt_family="LIPAl6"
31815..31838
/rpt_family="AT-rich"
32875..32910
/rpt_family="AT-rich"
32979..33163
/rpt_family="MLT1A1"
33164..33465
/rpt_family="AluJb"
33466..33616
/rpt_family="MLT1A1"
33704..34140
/rpt_family="MLT1K"
34745..34829
/rpt_family="MLT1K"
35312..35700
/rpt_family="L2"
36237..37252
/rpt_family="L1MC1"
complement(37557..38067)
/rpt_family="LTR40a"
39342..39757
/rpt_family="MLT2A2"
39760..39818
/rpt_family="(CATATA)n"
39819..39874
/rpt_family="MLT2A2"

Query Match 6.2%; Score 61.8; DB 9; Length 146883;
Best Local Similarity 57.5%; Pred. No. 0.00035;
Matches 111; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 796 AGGAGTTTGGGGGTTCCAGGATAGGAGTGGAGGTCAGAGGACGCAAGCAGCAGCA 855
Db 100174 AGGAGTGGAGGGGGGAGAAATGGGAGATGAGGTGAGGATACCAATA 100115

QY 856 TGTAGATGAACCGTCCAGAGAGCAAGCAGGAGAGCTGCAGGCCATCAGGCTGCA 915
Db 100114 TGTAAATAACAACCTCTGGAGACCCCTATGTACACCATGAGGACTATGTTTAAGAT 100055

QY 916 CTGTTCGTATTGGAGTTCATGCAAAATGAGTGTGTTTAGTCTCTTGCACAAAAA 975
Db 100054 TGTACTGTATTGGAGTTCTGCTAAATGAGTATATTTTAGTCTCTTGCACAAAAA 99995

QY 976 AAAAAAAAAAAAA 988
Db 99994 AAAAAATAAAAAA 99982

RESULT 24
AP429315 125020 bp DNA linear PRI 18-JAN-2002
LOCUS Homo sapiens junctophilin 3 (JPH3) gene, partial cds.
DEFINITION AP429315
ACCESSION AP429315
VERSION AP429315.1 GI:17646244
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 125020)
AUTHORS Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S.,
Ingersoll-Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A.,
Potter,N.T., Ross,C.A. and Margolis,R.L.
TITLE A repeat expansion in the gene encoding junctophilin-3 is
associated with Huntington disease-like 2
Nat. Genet. 29 (4), 377-378 (2001)
JOURNAL 21583737
MEDLINE 11694876
PUBMED 11694876
REFERENCE 2 (bases 1 to 125020)
AUTHORS Holmes,S.E., Ingersoll-Ashworth,R.G., Ross,C.A. and Margolis,R.L.

```

[illegible]

COMMENT

Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Apr 11, 2000 this sequence version replaced g1:7280308.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L4198

Center clone name: 276.N.1

----- Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Assembly: Dye-terminator Big Dye; 100% of reads

Consensus quality: Phrap; version 0.960731

Consensus quality: 148809 bases at least Q40

Consensus quality: 157187 bases at least Q30

Consensus quality: 160434 bases at least Q20

Insert size: 168000; agarose-fp

Insert size: 162462; sum-of-contigs

Quality coverage: 3.6 in Q20 bases; agarose-fp

Quality coverage: 3.7 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 24 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

1 1218: contig of 1218 bp in length
1219 1318: gap of 100 bp
1319 2383: contig of 1065 bp in length
2384 2483: gap of 100 bp
2484 3871: contig of 1388 bp in length
3872 3971: gap of 100 bp
3972 4836: contig of 865 bp in length
4837 4937: gap of 100 bp
4938 6262: contig of 1325 bp in length
6263 6362: gap of 100 bp
6363 8433: contig of 1971 bp in length
8434 11370: contig of 2938 bp in length
11371 11470: gap of 100 bp
11471 16136: contig of 4666 bp in length
16137 20360: contig of 4227 bp in length
20361 21060: gap of 100 bp
21061 25287: contig of 4227 bp in length
25288 32039: contig of 6652 bp in length
32040 32139: gap of 100 bp
32140 38768: contig of 6629 bp in length
38769 38868: gap of 100 bp
38869 43784: contig of 4916 bp in length
43785 50401: contig of 6517 bp in length
50402 57082: contig of 6581 bp in length
57083 57183: contig of 6735 bp in length
57184 64017: gap of 100 bp
64018 71454: contig of 7437 bp in length
71455 71554: gap of 100 bp
71555 79670: contig of 8116 bp in length
79671 79770: gap of 100 bp
79771 87413: contig of 7643 bp in length
87414 93773: gap of 100 bp
93774 93873: contig of 6260 bp in length
93874 93873: gap of 100 bp

```

```

* 93874 103590: contig of 9717 bp in length
* 103591 103690: gap of 100 bp
* 103691 118309: contig of 14619 bp in length
* 118310 118409: gap of 100 bp
* 118410 136058: contig of 17649 bp in length
* 136059 136158: gap of 100 bp
* 136159 164762: contig of 28604 bp in length.
* 136160 164762: contig of 28604 bp in length.

FEATURES
      Location/Qualifiers
      source
      1..164762
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /chromosome="9"
      /map="9"
      /clone="RP11-276N1"
      /clone_lib="RPC1-11 Human Male BAC"
      1..1218
      /note="assembly_fragment"
      1319..2383
      /note="assembly_fragment"
      2484..3871
      /note="assembly_fragment"
      3972..4836
      /note="assembly_fragment"
      clone_end:SP6
      vector_side:right"
      4937..6261
      /note="assembly_fragment"
      6362..8332
      /note="assembly_fragment"
      8433..11370
      /note="assembly_fragment"
      11471..16136
      /note="assembly_fragment"
      16237..20960
      /note="assembly_fragment"
      21061..25287
      /note="assembly_fragment"
      25388..32039
      /note="assembly_fragment"
      32140..38768
      /note="assembly_fragment"
      38869..43784
      /note="assembly_fragment"
      43885..50401
      /note="assembly_fragment"
      50502..57082
      /note="assembly_fragment"
      57183..63917
      /note="assembly_fragment"
      64018..71454
      /note="assembly_fragment"
      71555..79670
      /note="assembly_fragment"
      79771..87413
      /note="assembly_fragment"
      87514..93773
      /note="assembly_fragment"
      clone_end:TV
      vector_side:left"
      93874..103590
      /note="assembly_fragment"
      103691..118309
      /note="assembly_fragment"
      118410..136058
      /note="assembly_fragment"
      136159..164762
      /note="assembly_fragment"
      2301 others

BASE COUNT 47298 a 33699 c 32793 g 48671 t
ORIGIN
Query Match          6.1%; Score 60.8; DB 2; Length 164762;
Best Local Similarity 64.1%; Fred. No. 0.0006;

```

Matches 109; Conservative 0; Mismatches 57; Indels 4; Gaps 1;

QY 814 GGATAGGGGAATGGGAGGTTCAGAGGACGCAAGACGACGCCATGTAGATGAACCGTCCA 873

Db 64641 GAAATGGGAGATGTAGGTTCAGAGGACACAAAGTAAACACATGTAGGTGAACAGTCT 64700

QY 874 GAGAGCCAAAGCAGCGGAGAGGACTGAGGCCATCAGCGTGCACCTGTTCTGATTTGGAGTT 933

Db 64701 AGAGATCAAAATGTACGTGAAGATTACAGATAATAAAATTGTACT---GTATTGTGATT 64756

QY 934 CATGCAAAATGAGTGTGTTTGTAGTCTCTTGGCCACAAAAAATAAAAA 983

Db 64757 CCTGCTAAATGAGTAGACTTTAGCTATTCTTGCCACAAAAATAACAAA 64806

RESULT 26

AL353764/c

LOCUS Human DNA sequence from clone RP11-440G5 on chromosome 9, complete sequence.

ACCESSION AL353764

VERSION AL353764.9 GI:14272263

KEYWORDS HTG.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE

AUTHORS Tracey,A.

TITLE Direct Submission

JOURNAL Submitted (01-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

requests: clonerequest@sanger.ac.uk

On May 31, 2001 this sequence version replaced gi:13990016.

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone, and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep

This sequence was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr9

RP11-440G5 is from the library RPCT-11.2 constructed by the group of Pieter de Jong. For further details see

http://www.chori.org/bacpac/home.htm

VECTOR: pBAC3.6

This sequence is the entire insert of clone RP11-440G5. The true left end of clone RP11-818 is at 105684 in this sequence. The true right end of clone RP11-446A5 is at 70776 in this sequence.

FEATURES

Location/Qualifiers

1..170898

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosomes="9"

/clone="RP11-440G5"

/clone_lib="RPCT-11.2"

BASE COUNT 48782 a 34836 c 35124 g 52156 t

ORIGIN

Query Match 6.1%; Score 60.8; DB 9; Length 170898;

Best Local Similarity 64.1%; Pred. No. 0.0006;

Matches 109; Conservative 0; Mismatches 57; Indels 4; Gaps 1;

QY 814 GGATAGGGGAATGGGAGGTTCAGAGGACGCAAGACGACGCCATGTAGATGAACCGTCCA 873

Db 69053 GAAATGGGAGATGTAGGTTCAGAGGACACAAAGTAAACACATGTAGGTGAACAGTCT 68994

QY 874 GAGAGCCAAAGCAGCGGAGAGGACTGAGGCCATCAGCGTGCACCTGTTCTGATTTGGAGTT 933

Db 68993 AGAGATCAAAATGTACGTGAAGATTACAGATAATAAAATTGTACT---GTATTGTGATT 68938

QY 934 CATGCAAAATGAGTGTGTTTGTAGTCTCTTGGCCACAAAAAATAAAAA 983

Db 68937 CCTGCTAAATGAGTAGACTTTAGCTATTCTTGCCACAAAAATAACAAA 68888

RESULT 27

AL353764

LOCUS Human DNA sequence from clone RP11-1N7 from 2, complete sequence.

ACCESSION AC104684

VERSION AC104684.3 GI:18464246

KEYWORDS HTG.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE

AUTHORS Belter,E., Meyer,R. and Creason,K.

TITLE The sequence of Homo sapiens BAC clone RP11-1N7

JOURNAL Unpublished (2001)

Medline 99063792

2 (bases 1 to 129770)

3 (bases 1 to 129770)

Waterston,R.H.

Submitted (19-DEC-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

4 (bases 1 to 129770)

Waterston,R.H.

Submitted (01-FEB-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

5 (bases 1 to 129770)

Waterston,R.

Submitted (21-FEB-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

On Feb 1, 2002 this sequence version replaced gi:18250104.

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc>

Contact: sapiens@wustl.edu

----- Summary Statistics -----

Center project name: H_NH0001N07

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate

chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, E., Tateno, M., Catanese, J.J., and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org> and coworkers at <http://www.resgen.com> or Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is XXcos-2052C30; the clone sequenced to the right is RP11-356M6, 2000 bp overlap. Actual start of this clone is at base position 1 of RP11-IN7.

Sequence derived from one plasmid subclone, base position 94919 to 94957.

Data from AC067919 and AC097644 was used to finish this clone, AC104684.

Polymorphisms have been identified between AC067919, AC079779 and AC104684.

The sequence of AC021641 has been incorporated into AC104684.

FEATURES

source	Location/Qualifiers
	1..129770
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="2"
	/map="2"
	/clone="RP11-IN7"
	/clone_lib="RPCI-11"
repeat_region	1..77
repeat_region	/rpt_family="L1"
repeat_region	78..340
repeat_region	/rpt_family="Alu"
repeat_region	341..668
repeat_region	/rpt_family="L1"
repeat_region	661..932
repeat_region	/rpt_family="L1"
repeat_region	944..1216
repeat_region	/rpt_family="Alu"
repeat_region	1231..1277
repeat_region	/rpt_family="(TAAA)n"
repeat_region	1260..1341
repeat_region	/rpt_family="L1"
repeat_region	1342..1566
repeat_region	/rpt_family="Alu"
repeat_region	1567..1624
repeat_region	/rpt_family="L1"
repeat_region	1615..1636
repeat_region	/rpt_family="AT_rich"
repeat_region	1651..1778
repeat_region	/rpt_family="Alu"
repeat_region	3395..3458
repeat_region	/rpt_family="Achoho"
repeat_region	3459..3749
repeat_region	/rpt_family="Alu"
repeat_region	3750..3801
repeat_region	/rpt_family="L1"
repeat_region	3808..3939
repeat_region	/rpt_family="MIR"
repeat_region	8004..8034
repeat_region	/rpt_family="AT_rich"
repeat_region	8128..8236
repeat_region	/rpt_family="CR1"
repeat_region	10147..10658
repeat_region	/rpt_family="MaLR"
repeat_region	11021..11493
repeat_region	/rpt_family="MaLR"
repeat_region	11851..12002
repeat_region	/rpt_family="ERV1"
repeat_region	12684..12975
repeat_region	/rpt_family="Alu"
repeat_region	13127..13247
repeat_region	/rpt_family="MaLR"
repeat_region	13583..13678
repeat_region	/rpt_family="ERV1"
repeat_region	13679..14042
repeat_region	/rpt_family="ERV1"
repeat_region	15284..15648
repeat_region	/rpt_family="MaLR"
repeat_region	17182..17546
repeat_region	/rpt_family="MaLR"
repeat_region	18734..18848
repeat_region	/rpt_family="MIR"
repeat_region	18851..19169
repeat_region	/rpt_family="ERV1"
repeat_region	19405..19541
repeat_region	/rpt_family="MIR"
repeat_region	19818..20175
repeat_region	/rpt_family="MaLR"
repeat_region	21436..21561
repeat_region	/rpt_family="L1"
repeat_region	21587..22109
repeat_region	/rpt_family="ERV1"
repeat_region	23724..23845
repeat_region	/rpt_family="MIR"
repeat_region	23925..23987
repeat_region	/rpt_family="MIR"
repeat_region	24193..24332
repeat_region	/rpt_family="L2"
repeat_region	24255..24371
repeat_region	/rpt_family="MIR"
repeat_region	25991..26298
repeat_region	/rpt_family="Alu"
repeat_region	26273..26299
repeat_region	/rpt_family="AT_rich"
repeat_region	26380..26415
repeat_region	/rpt_family="AT_rich"
repeat_region	26417..26752
repeat_region	/rpt_family="L1"
repeat_region	26754..27066
repeat_region	/rpt_family="L1"
repeat_region	27067..27349
repeat_region	/rpt_family="Alu"
repeat_region	27350..27667
repeat_region	/rpt_family="L1"
repeat_region	27457..27494
repeat_region	/rpt_family="(CA)n"
repeat_region	28171..28277
repeat_region	/rpt_family="(TA)n"
repeat_region	28336..28384
repeat_region	/rpt_family="AT_rich"
repeat_region	28379..28554
repeat_region	/rpt_family="L1"
repeat_region	28916..29170
repeat_region	/rpt_family="L1"
repeat_region	29189..29500
repeat_region	/rpt_family="L1"
repeat_region	29511..29749
repeat_region	/rpt_family="L1"

```

repeat_region 29751..30738

Query Match
Best Local Similarity 6.1%; Score 60.4; DB 9; Length 129770;
Matches 112; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 794 AGAGGAGTTGGGGGTTTCAGATAGGAATGGGAGGTACAGAGCCAAAGCAGCAGC 853
Db 77435 AGGGGAGCTGGAGAGGGTGGGGGATGGAAGATATAGTCAAGAGGTACAAAGTAGCAGG 77494
QY 854 CATGTAGATCAACCGTCCAGAGAGCCAAAGCAGCAGCACTCCAGGCCATCAGCGTG 913
Db 77495 TGTGTGGATTAACAGTCCAGAGATCTCATGTACAATGAGACATAGTGTAAAAA 77554
QY 914 CACTGTTCTGATTTGGAGTTCATGCAAAATGAGTGTGTTTGTAGTCTCTTCCACAAAA 973
Db 77555 ATTGTGCTGTATTGGATTCTGATAATGAGTGTGATTGTAGTTGCTATTACCACAAA 77614
QY 974 AAAAAAARAAAAA 991
Db 77615 CAAAAAARAAAAA 77632

RESULT 28
AC067919/c AC067919 171655 bp DNA linear HTG 26-MAY-2000
DEFINITION Homo sapiens clone RP11-29216, WORKING DRAFT SEQUENCE, 20 unordered
pieces.
ACCESSION AC067919
VERSION AC067919.2 GI:8076830
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 171655)
Biren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-29216
Unpublished
2 (bases 1 to 171655)
Biren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Bonouslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collins,M., Cooke,P., DeAurellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Graud-Pierre,N., Grant,G., Hages,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Melidram,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tessaye,S., Theodore,J., Tirrell,A., Travers,M., Triglio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (27-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 25, 2000 this sequence version replaced gi:7652358.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIGR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

```

```

----- Project Information
Center project name: I6333
Center clone name: 292 I 6
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Assembly: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 161528 bases at least Q40
Consensus quality: 167303 bases at least Q30
Consensus quality: 168939 bases at least Q20
Insert size: 173000; agarose-fp
Insert size: 169755; sum-of-contigs
Quality coverage: 4.1 in Q20 bases; agarose-fp
Quality coverage: 4.1 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 220: contig of 220 bp in length
* 221 320: gap of 100 bp
* 321 4001: contig of 3681 bp in length
* 4002 4101: gap of 100 bp
* 4102 7945: contig of 3844 bp in length
* 7946 8045: gap of 100 bp
* 8046 11450: contig of 3405 bp in length
* 11451 11550: gap of 100 bp
* 11551 16117: contig of 4567 bp in length
* 16118 16217: gap of 100 bp
* 16218 20480: contig of 4263 bp in length
* 20481 20580: gap of 100 bp
* 20581 25760: contig of 5180 bp in length
* 25761 25860: gap of 100 bp
* 25861 31039: contig of 5179 bp in length
* 31040 31139: gap of 100 bp
* 31140 35820: contig of 4681 bp in length
* 35821 35920: gap of 100 bp
* 35921 42490: contig of 6570 bp in length
* 42491 42590: gap of 100 bp
* 42591 49780: contig of 7190 bp in length
* 49781 49880: gap of 100 bp
* 49881 57796: contig of 7916 bp in length
* 57797 57896: gap of 100 bp
* 57897 67541: contig of 9645 bp in length
* 67542 67641: gap of 100 bp
* 67642 76045: contig of 8404 bp in length
* 76046 76145: gap of 100 bp
* 76146 85834: contig of 9689 bp in length
* 85835 85934: gap of 100 bp
* 85935 95701: contig of 9767 bp in length
* 95702 95801: gap of 100 bp
* 95802 107607: contig of 11806 bp in length
* 107608 107707: gap of 100 bp
* 107708 123352: contig of 15645 bp in length
* 123353 123452: gap of 100 bp
* 123453 141544: contig of 18092 bp in length
* 141545 141644: gap of 100 bp
* 141645 171655: contig of 30011 bp in length.
----- Location/Qualifiers
1. 171655
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="RP11-29216"
/clone_lib="RPCI-11 Human Male BAC"
/notes="assembly_fragment
clone end:T7
vector_side:left"
----- misc_feature
1. 220

```

```

misc_feature 321..4001
    /note="assembly_fragment"
misc_feature 4102..7945
    /note="assembly_fragment"
misc_feature 8046..11450
    /note="assembly_fragment"
misc_feature 11551..16117
    /note="assembly_fragment"
misc_feature 16218..20480
    /note="assembly_fragment"
misc_feature 20581..25760
    /note="assembly_fragment"
misc_feature 25861..31039
    /note="assembly_fragment"
misc_feature 31140..35820
    /note="assembly_fragment"
    clone_end:SP6
vector_side:right"
misc_feature 35921..42490
    /note="assembly_fragment"
misc_feature 42591..49780
    /note="assembly_fragment"
misc_feature 49881..57796
    /note="assembly_fragment"
misc_feature 57897..67541
    /note="assembly_fragment"
misc_feature 67642..76045
    /note="assembly_fragment"
misc_feature 76146..85834
    /note="assembly_fragment"
misc_feature 85935..95701
    /note="assembly_fragment"
misc_feature 95802..107607
    /note="assembly_fragment"
misc_feature 107708..123352
    /note="assembly_fragment"
misc_feature 123453..141544
    /note="assembly_fragment"
misc_feature 141645..171685
    /note="assembly_fragment"
BASE COUNT 50815 a 35321 c 34768 g 48847 t 1904 others
ORIGIN
Query Match 6.1%; Score 60.4; DB 2; Length 171655;
Best Local Similarity 56.6%; Pred. No. 0.00075;
Matches 112; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
QY 794 AGAGAGTTTCGGGGGTTTCAGATAGGAAATGGGAGGTCAGAGGACGCCAAAGCAGCAGC 853
Db 150859 AGGGGAGCTGGAGAGGGTGGGGGATAGGTGGAAGTATAGGTCAGAAGTCAAAAGTAGCAGG 150800
QY 854 CATCTAGAACTCAACCGTCCAGAGACCCAGCAGCGGACGAGGACTGAGGCCATCAGCGTG 913
Db 150799 TGUGTGGATAAACAAGTCCAGAGATCTCATGTACAACATGAGGACTAGTGTAAAAAAA 150740
QY 914 CACTGTTCGTATTTGGAGTTTCATGCAAAATAGTGTGTTTGTAGCTGCTCTCCACAAA 973
Db 150739 ATTGTGCTGATTTGGATTCTGTATAATGAGTTGATTAGTTGCTATTACCACAAA 150680
QY 974 AAAAAAIAAAAAAAAAA 991
Db 150679 CAAAAACAAAAACAAA 150662

RESULT 29
AC010072
LOCUS AC010072 124347 bp DNA linear PRI 19-NOV-1999
DEFINITION Homo sapiens chromosome 14q31 clone CTD-217314 containing TSHR
gene, partial cds; and unknown gene, complete sequence.
ACCESSION AC010072
VERSION AC010072.5 GI:6453843
KEYWORDS HTG.
SOURCE Homo sapiens (human)

```


estimates computed by the Phrap assembly program.
 All manually edited bases have been reduced to quality zero.
 Quality levels above 40 are expected to have less than
 1 error in 10,000 bp.
 Base-by-base quality values are not generally visible from the
 GenBank flat file format but are available as part
 of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted:
 all regions were either double-stranded or sequenced with an
 alternate chemistry or covered by high quality data (i.e., Phred
 quality >= 30); an attempt was made to resolve all sequencing
 problems, such as compressions and repeats; all regions were
 covered by at least one plasmid subclone or more than one M13
 subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:

This sequence has been validated by Multiple Complete Digest
 fingerprinting. Comparison of the experimentally derived digest
 fragments with sequence-predicted fragments is given below.
 The electronically-digested sequence consists of both insert and
 vector, in order to accurately represent the entire circular BAC.
 Small fragments below a variable cutoff (approximately 400-800 bp)
 are not resolved in the fingerprint and hence do not appear
 in the table. There are no significant remaining discrepancies
 between the experimental and predicted values. Uniquely ordered
 fragments are separated by dashed lines.

BcoRI

HindIII

BglII

SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt
8696	8949	4503	4596	6422	6405
6	<800	6382	6584	2067	2089
6912	6827	512	<800	6249	6405
2142	2138	449	<800	470	<800
1633	1601	2282	2301	3028	2892
185	<800	404	<800	10089	9902
33	<800	3748	3800	2765	2775
235	<800	882	868	1438	1439
986	959	496	<800	5480	5750
5804	5786	1453	1429	8688	8596
166	<800	207	<800	2063	2089
2553	2584	4653	4596	7872	7784
73	<800	5302	5418	2057	2089
8410	8322	2615	2732	1970	2089
1588	1601	423	<800	344	<800
1392	1382	156	<800	5861	5750
3402	3455	3436	3385	965	988
147	<800	601	<800	4108	3983
1254	1250	52	<800	3172	3167
973	959	2019	2022	1145	1094

161	<800	3795	3800	1593	1595
313	<800	1540	1503	5146	5297
3713	3718	10075	9954	3644	3587
2627	2584	5447	5418	2365	2381
2804	2824	3913	3800	6608	6405
782	806	2683	2867	3181	3167
305	<800	622	<800	1074	988
2256	2264	1811	1749	1702	1736
4178	4181	1519	1503	759	<800
964	959	2068	2022	9566	9902
1486	1472	3072	3098	1419	1439
3586	3455	84	<800	9853	9902
6519	6434	381	<800	783	788
1447	1472	9	<800	4739	4629
2318	2264	534	<800	238	<800
1880	1840	388	<800	529	<800
1241	1250	2365	2540	2875	2892
3064	3099	5565	5418	205	<800
3415	3455	5400	5418	1223	1144
4194	4181	8055	8069	2821	2892
2586	2584	2087	2022	238	<800
2823	2824	857	868	939	988
3419	3455	613	<800	3195	3167
1501	1472	837	868	5065	4986
20	<800	3075	3216	1237	1247
4338	4390	6244	6292	5864	5750
5735	5786	3337	3385	1092	1094
4197	4181	2253	2301	379	<800
10513	10499	2916	2867	5118	4986
907	959	277	<800	989	988
842	806	1915	1912		
905	959	541	<800		
78	<800	574	<800		
12446	12330	7293	7311		
1636	1601	9275	9237		
1342	1382	2546	2622		
2807	2824	806	868		

Query Match	5.9%;	Score 58.6;	DB 9;	Length 151993;	COMMENT
Best Local Similarity	61.0%;	Pred. No. 0.0021;			
Matches 114;	Conservative 0;	Mismatches 69;	Indels 4;	Gaps 1;	
QY 797	GGAGTTTGGGGGTTTCAGATAGGGAATGGGAGGTTCAGAGGACCAAGCAGCAGCCAT	856			
Db 36052	GGATGTGGGGAGGGGAGGGAATGAGGAGATGTAGGTTCAGAGAATACAAAGTGGCAATAT	35993			
QY 857	GTAGAATGAACCGTCCAGAGAGCCAGCAGCGGACGAGCTCAGGCCATCAGCTGCAC	916			
Db 35992	GTATGATTACAGGTCAGATCTTAATATACCAATGAGATTAAGTAATAAATTGAC	35933			
QY 917	TCTTCGTTATTTGGAGTTCATGCAAAATAGTGTGTTTGTAGCTGCTTTGCCACAAAAA	976			
Db 35932	T---GTATGTGAGATTTCATGCTAAATAGTAGATTTTACCTGCTTTGCCAGGAACA	35877			
QY 977	AAAAAAA 983				
Db 35876	AACAGAA 35870				
RESULT 31					
AC012087					
LOCUS	AC012087	194635 bp	DNA	linear	PRI 18-AUG-2000
DEFINITION	Homo sapiens 3 BAC RP11-317L10, RP11-150K6, RP11-68D12 (Roswell Park Cancer Institute Human BAC Library) complete sequence.				
ACCESSION	AC012087 AC011608 AC012034 AC012016				
VERSION	AC012087.10 GI:7114465				
KEYWORDS	HTG.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Muzny, D.M., Adams, C., Bailey, M., Barbaria, J., Blankenburg, K., Bodota, B., Bouck, J., Bowie, S., Brooks, A., Buhay, C., Bunac, C., Burkett, C., Burrows, J., Carter, M., Chacko, J., Chen, Z., Cox, C., David, R., Delgado, O., Deshazo, D., Ding, Y., Domah-Rashid, N., Dugan-Rocha, S., Durbin, K.J., Fernandez, C., Ferraguto, D., Forcum-Tansey, J., Frantz, P., Ganesh, R., Garcia, D.K., Gorrell, J.H., Gorrell, L.L., Guevara, W., Harris, K., He, X., Hernandez, J., Hodgson, A., Hognes, M., Holloway, C., Hosak, H., Jackson, L.E., Jackson, L., Jia, Y., Jones, M., Kelly, S., Konderjewski, N., Kong, Y., Kovar, C., Leal, B., Li, Z., Lichtarge, O., Liu, J., Liu, W., Logan, O., Loez, R.J., Lu, J., Lucier, R., Martin, R., Martinez, C., McLeod, M.P., Mei, G., Moore, S., Moorish, T., Morgan, M., Morris, S., Nash, S., Nelson, A., Nguyen, R., Nguyen, N., Nguyen, S., Oswal, G., Parish, B., Paxton, S., Payton, B., Perez, L., Pu, L.L., Quiles, M., Reiter, D., Rives, M., Samuel, S., Say, J., Scherer, S., Shah, E., Shen, H., Simon, M., Sparks, A., Stamps, A., Suckang, R., Tabor, P., Taylor, T., Vasquez, L., Vinson, K., Vo, Q., Waibah, M., Watlington, S., Weinstock, G., Weinstein, I.R., Williamson, A., Worley, K., Wren, J., Wrenford, G., Yu, W., Zhou, X., Naylor, S.L., Nelson, D. and Gibbs, R.				
TITLE	Direct Submission				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 194635)				
AUTHORS	Worley, K.C.				
TITLE	Direct Submission				
JOURNAL	Submitted (20-OCT-1999) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA				
REFERENCE	3 (bases 1 to 194635)				
AUTHORS	Worley, K.C.				
TITLE	Direct Submission				
JOURNAL	Submitted (01-MAR-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA				
REFERENCE	4 (bases 1 to 194635)				
AUTHORS	Worley, K.C.				
TITLE	Direct Submission				
JOURNAL	Submitted (18-AUG-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One				

Baylor Plaza, Houston, TX 77030, USA
On or before Aug 18, 2000 this sequence version replaced
GI:6067112, GI:7025646, GI:7008674, GI:7025648.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3399-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT-----

Contig length: 194635
Phrap values in estimate: 193865
Average error rate (BOW-Phrap estimate): 3.31225e-05
Fraction of Phrap values less than 40 : 0.00865035
Number of consensus changing edits: 46
Number of N's in consensus : 0

----- Consensus changing edits -----
Position Original+Context Edited+Context
4981 agatggagat(n)tgatgggtg agatggagat(g)tgatgggtg
18700 agaactttta(n)ccaccagcc agaactttta(a)ccaccagcc
23735 tctaggagcc(n)taataagagt tctaggagcc(c)taataagagt
44381 ctgaatttga(n)ttcatataa ctgaatttga(a)ttcatataa
53445 aaaaaaaa(n)tcattgcaa aaaaaaaa(a)tcattgcaa
61587 tgcataaag(c)agatgggattc tgcataaag(c)agatgggattc
89906 tccaatgggt(g)tatgtgagt tccaatgggt(t)tatgtgagt
97122 caggccctt(c)ccccctct caggccctt(a)ccccctct
97318 agcccttgc(c)gtccccggg agcccttgc(b)gtccccggg
97464 tgcctgaag(c)tccttgagcc tgcctgaag(c)tccttgagcc
97608 aagaacagg(a)gggtggggcc aagaacagg(c)gggtggggcc
98041 gctttgctt(n)gggtgagga gctttgctt(a)gggtgagga
101579 caatctgctg(n)agatgtgtg caatctgctg(t)agatgtgtg
106238 aagcagtatg(n)cttcatctt aagcagtatg(a)cttcatctt
106254 tcttttttg(n)ttttttttt tcttttttg(a)ttttttttt
108707 ataagcattg(c)ctttgacct ataagcattg(t)ctttgacct
114218 gtaaatatg(g)taaatgaaa gtaaatatg(a)taaatgaaa
123681 gcgatggctg(c)ggnungggc gcgatggctg(g)ggnungggc
123682 cgatggctg(c)ggnungggc cgatggctg(c)ggnungggc

```

123684 atggctgcmg(n)ngmgcgcg
123685 tggctgcmg(n)ngmgcgcg
123686 gctgcmg(n)ngmgcgcg
123687 gctgcmg(n)ngmgcgcg
123688 ctgcmg(n)ngmgcgcg
123702 ggggtgcmg(n)ngmgcgcg
123720 ccttctctc(n)ctgcccct
123731 ctgcccct(n)gagagcgg
123763 ctggcgcc(n)cgccaccg
123680 taggaagga(n)tgctctggac
137763 aaaaaaaa(n)ccagaagaa
137764 aaaaaaaa(n)ccagaagaa
148155 gagtcagga(n)ttcgagaca
151494 ttctcaga(n)atcattct
163884 aaaaaaaa(n)aaagtaagtc
171028 tactcataa(n)taattaccac
181695 taggaacatg(n)aatattgtc
185127 agctggaac(n)tcagcctct
185131 ggaattct(n)acnccatcc
185133 attcctnact(n)ccatccagg
185153 ctttcccca(n)gtaattatg
185191 cactttatga(n)gnaataaam
185193 ccttatgag(n)aaataunga
185200 angnaataa(n)naagggcagt
185201 ngnaataa(n)gagggcagt
190487 acaacagca(n)caactgctgtg
190499 actgtggt(n)caactggaag
192055 aagcctgtg(n)ttctagacgt

```

----- Distribution of Quality < 40 Bases -----

# bases	5	10	15	20	25	30	35	40
1000								
800								
700								
600								
500								
400								
300								
200								
100								
0								

FEATURES

source

Version: 1.01 qxf.

```

Location/Qualifiers
1. .194635
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-317L10, RP11-150K6, RP11-68D12"
complement(40. .496)
/rpt_family="L1M4"
1423. .1452
/rpt_family="AT_rich"
1705. .2095
/rpt_family="L1R16A"
complement(2399. .2430)
/rpt_family="L2"
complement(2431. .2954)
/rpt_family="MLT1F"
complement(2955. .3056)
/rpt_family="L2"
complement(3109. .3199)
/rpt_family="L2"
5566. .5781
/rpt_family="MER20"
6399. .6418

```

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

```

repeat_region
repeat_region
repeat_region

```

Query Match 5.9%; Score 58.6; DB 9; Length 194635;
Best Local Similarity 61.0%; Pred. No. 0.002;
Matches 114; Conservative 0; Mismatches 69; Indels 4; Gaps 1;

```

QY 797 GGAGTTTGGGGGTCAGGATAGGAATGGGAGGTCAGAGGAGCGAAACGACGACCAT 856
Db 54877 GGATGTGGGAGGCGAGGAATGAGGATGTTAGTCAGAGATACAAATGGCAATAT 54936
QY 857 GTAGATGAACCTCCAGAGAGCAACGAGGAGGAGGATGCGAGGCGATCAGGTCGAC 916
Db 54937 GTATGATTACAGGTGAGATCTAATATACAAATGAGATTAATAGGTAATAAATTTGAC 54996
QY 917 TGTTCTGTTTGGAGTTCATGCAAAATGAGTGTGTTTAGTCTCTTGGCCACAAAAA 976
Db 54997 T-----GTATGTGAGATTCATGCTAAATGAGTAGATTTTACCTCTCTTGGCCACAAAAA 55052
QY 977 AAAAAA 983
Db 55053 AACAGAA 55059

```

RESULT 32

HSML30AC2/c

```

LOCUS HSML30AC2 4950 bp mRNA linear PRI 22-NOV-1996
DEFINITION H.sapiens mRNA for M130 antigen cytoplasmic variant 2.
ACCESSION Z22970
VERSION Z22970.1 GI:312145
KEYWORDS antigen; antigen M130.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4950)
AUTHORS Law S.K., Micklem, K.J., Shaw, J.M., Zhang, X.P., Dong, Y., Willis, A.C.
and Mason, D.Y.
TITLE A new macrophage differentiation antigen which is a member of the
scavenger receptor superfamily
Eur. J. Immunol. 23 (9), 2320-2325 (1993)
JOURNAL 93380506
MEDLINE 8370408
PUBMED
REFERENCE 2 (bases 1 to 4950)
AUTHORS Micklem, K.K.
DIRECT SUBMISSION
TITLE Submitted (14-JUN-1993) Kingsley K.J. Micklem, Nuffield Department
of Pathology, University of Oxford, Level 1, Maternity Block, John
Radcliffe, Hospital, Headington, Oxford, OX3 9DU, United Kingdom
LOCATION/QUALIFIERS
1. .4950
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_type="Macrophage"
102. .3572
/codon_start=1
/product="M130 antigen cytoplasmic variant 2"
/protein_id="CAA80543.1"
/db_xref="GI:312146"
/db_xref="SPTREMBL:Q07900"
/translation="MWLEDSGSADFRHFVNLSPFTITVLLSACFVTSLSGDTKD
ELRLVDGNGKSGRVEVKVQBEWTCVNGNGMEAVSVICNLQCGTAKPAGWANS
AGSGRIWMDHVSRCRGNLSALWDCKDGNKHSNCTHOODAGVTCDSGNLWRLTRGG
NMCSGRLEIKFQGRWTVCDNFNIDHSAVICRQLEGSVSVFSGSNFGSGPIWF
DDLICNGESALMNCKHQGKHNCKHAEDAGVTCRSGADLSRLDVGVTCSGRLEV
RFQGEWGTICDDGWDSDYAAVACKQCGPTAVTAIGRVNASKGFHGLWLDVSQGHV
PAWQCKHEWKGHYCNHNEADGVTCSGDLRLRGGSRCACTVEVETLRLGKV
CDRGWGLKADVVCRLQCGSALTKSYQVSKIQTNTWLFSLSCNGNETSLWCKKNW
QWGLGLTCDHYEEAKITCSAHPRLVGDIPCSGRVEVKHGDITWGSICDSFSLAAS

```

CDS

```

VLCRELCQCTVSVILGGAHGRNGOIAWEEFQCECHESHLSICPVAPRPEGTCTSHSR
DVGVCSSRYTEIRLVNKGTPCGRVELKTLGWSLNCNHWIDIEAHVLCQOLKCGVA
LSTPGARGKNGQITPIEFHMFCTQEQMGDCPVTALGASLCPSEVAEIVSGTCSGNQS
QTLSCNSSLGCTRTPIEESAIVACIESQGLRLVNGGRCAGREIIEHSGVGTICD
DSWDLSDHVGCRQIGCEAINATSAHFEGETPTLWDEKMGKESRWCHSHGV
GQONCRHKEDAGVICSEFMSLRITSEAREACAGLEVDYFNGAWGTGKSGSETTVG
VVCROLGADKGIKIPASLDKAMSIPMWVNVQCPKPDPTLWQCPSSPWEKELASPSE
EWITCDNKIRIQEGPTSCSREVEIWHGSGSWCTVQDSDWDLDDAQVVOQLGCGPALK
APKEAFEGQGTPIWNEVKCKNGESSLWDCPARRWGHSECKHDEAAVNCVTDISVQK
TPQKATIGRSSKQSEFLAVIGLGVLLAIFVALFTLKRRQRQRLAYSSRGENLVHQ
IYREMNCSCLNADDDLLMNSGLVLGGSIAQFRSVAAVEAQTFFYFDKQKKXNVI
GSLDAYNGQE"
102. .221
222. .3569
/product="M130 antigen cytoplasmic variant 2"
/evidence=experimental
4924. .4929
polyA_signal
4928. .4933
polyA_signal
4932. .4937
polyA_signal
4947
BASE COUNT 1348 a 970 c 1277 g 1355 t
ORIGIN
Query Match 5.9%; Score 58.4; DB 9; Length 4950;
Best Local Similarity 59.0%; Pred. No. 0.0025;
Matches 118; Conservative 0; Mismatches 81; Indels 1; Gaps 1;
QY 787 TGAGTGAAGAGAGTTGGGGGTTGAGGATAGGAATGGGGAGTCTAGAGACCAAG 846
Db 4080 TGACAGGAATGGGAATGAAGAACTGTGAAGACGCTATGTAGTCTAGAGGATACAAAG 4021
QY 847 CAGCAGCCATGATGAAATGAACCTGCAGAGAGCCAGCAGCAGGACTGCAGGCCAT 906
Db 4020 TAGCAGATACGTAGATGACACGCTAGAGCTCTAATGTACAACTAGAGATTATAGT 3961
QY 907 CAGCGTGCACTGTCGTATTGGAGTTTCATGCAGAAATCAGTGTGTTTGTGCTCTTTC 966
Db 3960 AATAACATTGTGCT-GTATGTAGGATTCATGCTAAGTGAAGTGTGTTTGTGCTCTTTC 3902
QY 967 CACAAAAAACAACAAAAA 986
Db 3901 CACAAGAAACAACAACAAAAA 3882
RESULT 33
AC131206 177260 bp DNA linear PRI 31-DEC-2002
LOCUS Homo sapiens 12 BAC RP11-157G21 (Roswell Park Cancer Institute
DEFINITION Human BAC Library) complete sequence.
ACCESSION AC131206
VERSION AC131206.2 GI:22549585
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 177260)
Muzny D.M., Adams C., Adio-Oduola B., Ali-osman, P.R., Allen, C.,
Albrooks, S.L., Amaratunge, H.C., Are, J.R., Ayale, M., Banks, T.,
Barbaria, J., Benton, J., Bimage, K., Blankenburg, K., Bomlin, D.,
Bouck, J., Bowie, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carroll, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chiu, D., Chowdhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
Dolaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Emerling, S.,
Escotto, M., Falla, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P.,
Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N.,
Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S.,
Hamilton, K., Han, J., Harris, C., Harris, K., Hart, M., Haviak, P.,

```

Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Ioshikhes, I., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Lead, B., Lee, B., Lewis, L., Lewis, L., Li, J., Li, Z., Lichtarge, O., Liu, C., Liu, J., Liu, W., Loulseghe, H., Lozado, R., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Marondel, I., Martin, R., Martindale, A., Martinez, E., Massey, E., Mathew, E., McLeod, M.P., Meador, M., Mei, G., Merscher, S., Mohabbat, K., Montgomery, K.T., Morgan, Z., Mitchell, I., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G., Oragunye, N., Oviado, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savory, G., Scherzer, S., Scott, G., Shen, H., Shim, C., Shoostari, N., Sisson, I., Sutton, A., Svatik, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Zucherlapati, R., Weinstock, G. and Gibbs, R.

Direct Submission

Unpublished

2 (bases 1 to 177260)

Worley, K.C.

Direct Submission

Submitted (18-AUG-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 177260)

Worley, K.C.

Direct Submission

Submitted (30-AUG-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

4 (bases 1 to 177260)

Worley, K.C.

Direct Submission

Submitted (25-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

5 (bases 1 to 177260)

Worley, K.C.

Direct Submission

Submitted (31-DEC-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Aug 30, 2002 this sequence version replaced gi:22296911.

INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email

gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence

continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.ht> ml.

FEATURES

source

Location/Qualifiers

1..177260

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="12"

/clone="RP11-157G21"

1..2001

/note="overlaps bases 44655..46655 of clone AC131207"

/function="clone overlap"

64444..64708

/standard name="D12S1890"

68193..68303

/standard name="D10S275"

78117..78366

/standard name="STS-H93331"

90052..90229

/standard name="D11S2442"

90092..90372

/standard name="D11S3206"

93532..93750

/standard name="D11S3059"

137355..137635

/standard name="D11S3206"

147420..147700

/standard name="D11S3206"

147563..147740

/standard name="D11S2442"

175236..177260

/note="overlaps bases 1..2025 of clone AC006927"

/function="clone overlap"

BASE COUNT 59476 a 33910 c 33289 g 50585 t

ORIGIN

Query Match 5.9%; Score 58.4; DB 9; Length 177260;

Best Local Similarity 59.0%; Pred. No. 0.0023;

Matches 118; Conservative 0; Mismatches 81; Indels 1; Gaps 1;

QY 787 TGAGTGAAGAGGAGTTGGGGGGTTTCAGGATAGGGAATGGGAGGTTCAGAGCAGCAAG 846

Db 64219 TGACAGGAATGGGAATGAAGAACTGTGAAGACGTATGTAGTTCAGAGATCAAG 64278

QY 847 CAGAGCCATGTAGATGAACCGTCCAGAGAGCCAGCAGCAGGAGCACTGCAGGCCAT 906

Db 64279 TAGCAGATACGTAGATGAACAGCTAGAGGTCTTAATCAACATCAGGATTATAGT 64338

QY 907 CAGCGTGACTTGTGTTATTCGAGTTCATGCAAAATGAGTGTGTTAGTGTCTTTCG 966

Db 64339 AATAACATTGTCT-GTATGTAGGATTCATGTAAGTAGTAGTGTGTTAGTGTCTTTCG 64397

QY 967 CACAAAAAATAAAAAA 986

Db 64398 CACAAGACACACAAAAA 64417

RESULT 34

AC0233003
LOCUS
DEFINITION
AC0233003
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
JOURNAL
TITLE
COMMENT

AC0233003
Homo sapiens clone RP11-270M20, WORKING DRAFT SEQUENCE, 12
unordered pieces.
AC0233003
HTG; HTGS PHASE1; HTGS_DRAFT.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 140055)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-270M20
Unpublished
2 (bases 1 to 140055)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F., Castle,A.,
Boguslavsky,L., Bouckgater,B., Brown,A., Burkett,G., Castle,A.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeArelano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lechoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheeters,R., Meldrum,J., Meneus,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Olivari,T.M., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (07-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141 USA
On Mar 1, 2000 this sequence version replaced gi:6921586.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIRK
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L6597
Center clone name: 270 M.20
----- Summary Statistics
Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 133511 bases at least Q40
Consensus quality: 136178 bases at least Q30
Consensus quality: 137450 bases at least Q20
Insert size: 139000; agarose-fp
Insert size: 138955; sum-of-contigs
Quality coverage: 5.1 in Q20 bases; agarose-fp
Quality coverage: 5.1 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1060: contig of 1060 bp in length
* 1061 1160: gap of 100 bp
* 1161 3767: contig of 2607 bp in length
* 3768 3867: gap of 100 bp
* 3868 8490: contig of 4623 bp in length

```

* 8491      8590: gap of 100 bp
* 8591      14297: contig of 5707 bp in length
* 14298      14397: gap of 100 bp
* 14398      20798: contig of 6401 bp in length
* 20799      20898: gap of 100 bp
* 20899      31912: contig of 11014 bp in length
* 31913      32012: gap of 100 bp
* 32013      43544: contig of 11532 bp in length
* 43545      43644: gap of 100 bp
* 43645      52029: contig of 8385 bp in length
* 52030      52129: gap of 100 bp
* 52130      67985: contig of 15856 bp in length
* 67986      68085: gap of 100 bp
* 68086      84647: contig of 16562 bp in length
* 84648      84747: gap of 100 bp
* 84748      108247: contig of 23500 bp in length
* 108248      108347: gap of 100 bp
* 108348      140055: contig of 31708 bp in length.

FEATURES
    source
    1..140055
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /clone_lib="RP11-270M20"
        /clone_lib="RPCI-11 Human Male BAC"
    misc_feature
    1..1060
        /note="assembly_fragment"
    misc_feature
    1161..3767
        /note="assembly_fragment"
    misc_feature
    3868..8490
        /note="assembly_fragment"
    misc_feature
    8591..14297
        /note="assembly_fragment"
    misc_feature
    14398..20798
        /note="assembly_fragment"
        clone_end:17
        vector_side:right
    misc_feature
    20899..31912
        /note="assembly_fragment"
    misc_feature
    32013..43544
        /note="assembly_fragment"
    misc_feature
    43645..52029
        /note="assembly_fragment"
        clone_end:86
        vector_side:left
    misc_feature
    52130..67985
        /note="assembly_fragment"
    misc_feature
    68086..84647
        /note="assembly_fragment"
    misc_feature
    84748..108247
        /note="assembly_fragment"
    misc_feature
    108348..140055
        /note="assembly_fragment"

BASE COUNT  42310 a 26393 c 26935 g 43312 t 1101 others
ORIGIN
Query Match      5.8%; Score 58; DB 2; Length 140055;
Best Local Similarity 61.9%; Pred. No. 0.0029;
Matches 125; Conservative 0; Mismatches 75; Indels 2; Gaps 2;
QY 787 TGAGTGAAGAGGAGTTTGGGGGTTTCAGGATAGGAGTGGGGAGTTCAGAGGAGGCAAG 846
Db 15860 TGACAGAGGAGGTAGAAAATGCATAGGAATGGGAATATATAGGTTCAGAGGACACAAAG 15919
QY 847 CAGCAGCCATGTAATGAACC-GTCCAGAGGCAAGCAAGCGGACGAGGACTGCAGGCGCA 905
Db 15920 CAGCAGCCATGAGGAGGACCAAGTCTAGAGATCTTAAGGTACATATGAGGACCATACAG 15979
QY 906 TCAGGTGCATCTGTTGATTTGAGTTCATGCAAAATGAGTGTGTTTACGTCCTTG 965
Db 15980 GTAAATAGAGGCGTACTGTGTATGGATTCATGCTAAACTAGTA-GAATTGGCTGCTCTTG 16038
QY 966 CCACAAAAAAAAAAAAAAAAA 987

```

Db 16039 CCACAAAAACAAACAAAAAA 16060

RESULT 35
ALI39342

LOCUS

DEFINITION

ALI39342 151553 bp DNA linear PRL 03-JAN-2002

Human DNA sequence from clone RP5-1016N21 on chromosome 1q42.13-43. Contains part of the gene for a novel protein (ortholog of Drosophila Pecanex (PCX), similar to KIAA0805), part of the gene for a novel protein similar to KIAA0995, STSs and GSSs, complete sequence.

ACCESSION

ALI39342

VERSION

ALI39342.7 GI:9796485

KEYWORDS

HTG; KIAA0805; KIAA0995; PCX; Pecanex.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 151553)

AUTHORS

Cobley, V.

TITLE

Direct Submission

JOURNAL

Submitted (05-DEC-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

COMMENT

requests: clonerequest@sanger.ac.uk

On Aug 12, 2000 this sequence version replaced gi:9712646.

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WormPEP; Information on the WormPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr1>

This sequence is the entire insert of clone RP5-1016N21. The true left end of clone RP11-109B17 is at 90706 in this sequence. The true right end of clone RP11-109B17 is at 77343 in this sequence. This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated repeat sequence elements. Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key. RP5-1016N21 is from the library RPCI-5 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pCYAC2.

FEATURES

location/Qualifiers

1..151553

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="1"

/map="q42.13-43"

/clone="RP5-1016N21"

/clone_lib="RPCI-5"

1..77

/note="Alu repeat: matches 232..310 of consensus"

101..408

/note="L1MC/D repeat: matches 5263..5571 of consensus"

465..619

/notes="L1MD2 repeat: matches 5810..5953 of consensus"

628..729

/notes="L1MEC repeat: matches 2110..2213 of consensus"

836..1468

/note="L1M4 repeat: matches 4267..4889 of consensus"

1511..1685

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

/note="L1MC3 repeat: matches 6433. .6606 of consensus"
1745. .2410
/note="L1MC3 repeat: matches 6609. .7281 of consensus"
2411. .2705
/note="AluSg repeat: matches 1. .296 of consensus"
2706. .3163
/note="L1MC3 repeat: matches 7281. .7736 of consensus"
3183. .3479
/note="AluSg repeat: matches 1. .294 of consensus"
complement (4350. .4820)
/note="match: GSS: Em:AQ827082"
4999. .5225
/note="L2 repeat: matches 1328. .1564 of consensus"
5811. .6105
/note="AluSx repeat: matches 1. .295 of consensus"
6180. .6371
/note="L1MC3 repeat: matches 5725. .5926 of consensus"
6386. .6421
/note="18 copies 2 mer gt 97% conserved"
6673. .7475
/note="L2 repeat: matches 1777. .2614 of consensus"
7485. .7709
/note="MIR repeat: matches 21. .262 of consensus"
7733. .7783
/note="L2 repeat: matches 2648. .2705 of consensus"
8206. .8517
/note="AluSx repeat: matches 1. .312 of consensus"
8833. .10173
/note="PTP5 repeat: matches 743. .2438 of consensus"
10349. .10385
/note="19 copies 2 mer tg 100% conserved"
10388. .10664
/note="AluSg repeat: matches 1. .306 of consensus"
complement (13145. .13540)
/note="match: GSS: Em:AQ177321"
13457. .13718
/note="AluJo repeat: matches 1. .258 of consensus"
14443. .14562
/note="6 copies 20 mer 78% conserved"
14515. .14558
/note="22 copies 2 mer ag 77% conserved"
complement (14896. .15073, 19597. .19738, 40166. .40316,
41154. .41246, 57685. .57781, 58949. .59058, 78822. .78913,
80037. .80147, 88401. .88572, 97782. .97814, 97914. .98067,
107109. .107254))
/genes="dJ1016N21.1"
complement (join(<14896. .15073, 19597. .19738, 40166. .40316,
41154. .41246, 57685. .57781, 58949. .59058, 78822. .78913,
80037. .80147, 88401. .88572, 97782. .97814, 97914. .98067,
107109. .107254))
/genes="dJ1016N21.1"
/note="match: CDNAS: Em:AF096286 Em:AB018348
match: proteins: Tr:Q9QYCI Tr:Q94897 Tr:Q9W507 Tr:O46074
Tr:Q9W506 Tr:O61825 Sw:PL8490"
/codon_start=1
/evidence=not_experimental
/product="dJ1016N21.1 (novel protein (ortholog of
Drosophila Pecanex (PCX), similar to KIAA0805))"
/protein_id="CAD20144.1"
/db_xref="GI:18070856"
/translation="HYSDLEASSCSSTOCKENREOFYKFIIPGKWIKWYDRLLTLL
ALLDREDKENVAILLLVLSLGLFGLSGFCCKMWWLLFLCLVMASCCQYSLKSV
QDPSPFIHGHQIITYPTVFCVLGLLIDTGAKARHPFVYVYGLKLFSPVFL
QSDVLYVLCXFAISLGLFPQINTFTLLLEQIDMLFFGGSAVSITSAVSVFA
RSVLAAILHVCFSALKPEPWSMOHLPALFSAFCGLVALSVHLSRQSDSDPSVLSFI
QCLRPFKLHONLAESAADPLPKMKDSDVDLKVLDLIVCAVAVLSFAVSAVSFVLS
LRPFLSIVLALAGAVGFVTHVLPQLRKHPHWMWLSHFIILKNKETHQREVRDVAHLM

WFERLYVWLOCFEKYLLYPALILNALTIDAFILSNHRLGTHWDIFLMIAGMKLLRT
SFQNPYQFINLSFTVIFDFDYKDISFSLDFPMVSIILFSK"
15184. .15344
/note="MIR repeat: matches 12. .179 of consensus"
16957. .17561
/note="MER6A repeat: matches 1. .598 of consensus"
complement (18841. .19304)
/genes="dJ1016N21.1"
/note="match: GSS: Em:AQ558297"
complement (18915. .19311)
/genes="dJ1016N21.1"
/note="match: GSS: Em:AQ338718"
20315. .20425
/note="MIR repeat: matches 18. .129 of consensus"
21268. .21570
/note="AluSx repeat: matches 1. .297 of consensus"
21571. .21690
/note="6 copies 20 mer 80% conserved"
22427. .23076
/note="L2 repeat: matches 1148. .1840 of consensus"
23241. .23347
/note="MER5A repeat: matches 4. .112 of consensus"
23351. .23520
/note="MER5A repeat: matches 26. .188 of consensus"
23676. .23858
/note="L2 repeat: matches 953. .1144 of consensus"
23868. .24080
/note="AluJb repeat: matches 87. .298 of consensus"
24270. .24836
/note="LIME repeat: matches 5092. .5651 of consensus"
24837. .25148
/note="AluY repeat: matches 1. .310 of consensus"
complement (24980. .25609)
/genes="dJ1016N21.1"
/note="match: GSS: Em:AQ344097"
25149. .25356
/note="LIME repeat: matches 4886. .5093 of consensus"
25566. .25747
/note="MER97b repeat: matches 1. .187 of consensus"
complement (25605. .26037)
/genes="dJ1016N21.1"
/note="match: GSS: Em:AQ683895"
25810. .25859
/note="trigger4 (Zombi) repeat: matches 1. .48 of consensus"
25860. .25907
/note="MER46A repeat: matches 49. .97 of consensus"
25900. .25979
/note="MER46A repeat: matches 153. .236 of consensus"
26937. .27418
/note="match: GSS: Em:AQ561984"
27022. .27459
/note="TIGGR2 repeat: matches 2280. .2716 of consensus"
27471. .27530
/note="MER8 repeat: matches 178. .239 of consensus"
28304. .28584
/note="AluSg repeat: matches 1. .310 of consensus"
29471. .29912
/note="match: GSS: Em:B66297"
29498. .29966
/note="match: GSS: Em:AQ229439"

Query Match 5.8%; Score 58; DB 9; Length 151553;
Best Local Similarity 59.7%; Pred. No. 0.0029;
Matches 117; Conservative 0; Mismatches 75; Indels 4; Gaps 1;
QY 789 ACTGAAGAGGAGTTTGGGGGTTTACGATAGGAATGGGAGGTGAGGAGCGAAGCA 848
DB 47275 AGTTAACTATGGCATGGGGAGGTAGAGATGGGATATGTGTGTAGAAATTTATAAGTA 47334
QY 849 GCAGCCGTGTAAGTGAACCGTCCAGAGAGCAACGACGAGGATGCGAGGCATCA 908
DB 47335 GCAGATATGTAGGATAAACAACCTCCAGCGATCTAATGTACATGAGAACTATAATTA 47394

```

QY 909 GCGTCACCTGTTGTTATTTGGAGTTCATGCAAAATGAGTGTGTTAGCTGCTTGGCA 968
Db 47395 AACTGTACT-----ATATTGGGATTCCTGCAAAATAGTAGATTTAGCTGCTTGGCA 47450

QY 969 CAAAAAATAAAAAA 984
Db 47451 CAAAAAATAAAAAA 47466

RESULT 36
AC019159/c
LOCUS AC019159 163085 bp DNA linear PRI 07-NOV-2001
DEFINITION Homo sapiens BAC clone RP11-56018 from 2, complete sequence.
ACCESSION AC019159
VERSION AC019159.8 GI:13677116
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Sulston,J.E. and Waterston,R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
99063792
9847074
PUBMED
2 (bases 1 to 163085)
Goyea,B., Cotton,M., Spalding,L. and Lehnert,L.
The sequence of Homo sapiens BAC clone RP11-56018
Unpublished
3 (bases 1 to 163085)
Waterston,R.H.
Direct Submisson
Submitted (30-DEC-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 163085)
Waterston,R.H.
Direct Submisson
Submitted (19-APR-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 163085)
Waterston,R.H.
Direct Submisson
Submitted (20-APR-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
6 (bases 1 to 163085)
Waterston,R.
Direct Submisson
Submitted (07-NOV-2001) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Apr 19, 2001 this sequence version replaced gi:11276269.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@watson.wustl.edu
----- Summary Statistics
----- Center project name: H_NH0056018

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence

from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RP11-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P. Y., Zhao, B., Frengen, E., Tatenio, M., Catanese, J. J., and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is AC027111; the clone sequenced to the right is AC023003. Actual start of this clone is at base position 1 of RP11-56018; actual end is at base position 163085 of RP11-56018.

The sequence from base position 822245 to 82358 is derived from a single plasmid subclone. Size information based on PCR supports the assembly.

FEATURES	source	Location/Qualifiers
	1..163085	
	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:9606"	
	/chromosome="2"	
	/map="2"	
	/clones="RP11-56018"	
	/clone_lib="RP11-11"	
	1..51	
	/note="similar to EST AV696391 (NID:gl0298254)"	
	107..224	
	/note="similar to EST AV696391 (NID:gl0298254)"	
	197..373	
	/rpt_family="MIR"	
	568..667	
	/rpt_family="MIR"	
	668..775	
	/rpt_family="MER1_type"	
	1249..1439	
	/rpt_family="MIR"	
	8077..8445	
	/rpt_family="MALR"	
	8451..9804	
	/rpt_family="MALR"	
	9861..10142	
	/rpt_family="Alu"	
	10149..10387	
	/rpt_family="MALR"	
	10393..10790	
	/rpt_family="MALR"	
	11489..11710	
	/rpt_family="MIR"	
	11985..12022	
	/rpt_family="(TC)n"	
	12131..12886	
	/rpt_family="L1"	
	13153..13392	
	/rpt_family="L1"	
	13397..13441	
	/rpt_family="(TC)n"	
	13513..13577	
	/rpt_family="(TA)n"	
	13578..14029	
	/rpt_family="L1"	

Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Totham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.-J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (01-APR-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Apr 1, 2002 this sequence version replaced gi:18643501.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L13055
 Center clone name: 2206_N_4

Only the first 99,2 kilobases of this clone are being submitted.
 The remainder of the clone is overlapped by either accession number
 AC006441 [WICGR project L515] or
 accession number AC004408 [WICGR project L309].
 Location/Qualifiers

FEATURES

```

source
1..99268
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /chromosome="17"
  /map="17"
  /clone="CTD-2206N4"
  /clone_lib="CITD1 Human BAC"
  283..1026
    /rpt_family="tigger1"
  1032..1324
    /rpt_family="AluSc"
  1325..1501
    /rpt_family="tigger1"
  complement(1502..1804)
    /rpt_family="AluSc"
  1805..2245
    /rpt_family="tigger1"
  2246..2567
    /rpt_family="tigger1"
  2708..2999
    /rpt_family="tigger1"
  3369..3396
    /rpt_family="AluSc"
  complement(3640..3935)
    /rpt_family="AluY"
  complement(3903..3908)
    /note="<30 qual SNGL region"
  complement(3915..3919)
    /note="<30 qual SNGL region"
  4083..4237
    /rpt_family="AluSc"
  complement(4938..5073)
    /rpt_family="AluSc"
  complement(5107..5413)
    /rpt_family="AluSc"
  complement(6121..6407)
    /rpt_family="AluSc"
  6553..6870
    /rpt_family="AluY"
  6562..6665
    /note="single clone coverage"
  6620..6656
    /note="<30 qual SNGL region"
  6733..6741
    /note="<30 qual SNGL region"
  6752..6761
    /note="<30 qual SNGL region"
  
```

```

  6776..6780
    /note="<30 qual SNGL region"
  6814..6824
    /note="<30 qual SNGL region"
  6887..7158
    /rpt_family="AluJo"
  7064..7069
    /note="<30 qual SNGL region"
  7161..7227
    /rpt_family="GGAA)n"
  7385..7474
    /rpt_family="AluSg/x"
  complement(7510..7804)
    /rpt_family="AluSp"
  7540..7551
    /note="<30 qual SNGL region"
  9096..9436
    /rpt_family="AluY"
  13978..14187
    /rpt_family="MIR"
  15639..15660
    /rpt_family="AT rich"
  complement(15685..15997)
    /rpt_family="AluJo"
  complement(16062..16730)
    /rpt_family="L1ME4A"
  complement(16940..17144)
    /rpt_family="AluSx"
  17854..18011
    /rpt_family="L1M4"
  complement(18099..18424)
    /rpt_family="MER3"
  18451..18471
    /rpt_family="AT rich"
  18492..18784
    /rpt_family="AluJo"
  complement(18788..18890)
    /rpt_family="MER3"
  complement(18894..19002)
    /rpt_family="MER3"
  19029..19140
    /rpt_family="L1ME3"
  19181..19229
    /rpt_family="(TTT)n"
  complement(19242..19488)
    /rpt_family="AluJo"
  complement(19497..19849)
    /rpt_family="AluSg"
  20030..20162
    /rpt_family="FLAM_C"
  20424..20478
    /rpt_family="(TG)n"
  20479..20500
    /rpt_family="(TGG)n"
  20791..20871
    /rpt_family="MIR3"
  21057..21245
    /rpt_family="MER58C"
  21249..21539
    /rpt_family="AluSx"
  21654..21707
    /rpt_family="AT rich"
  complement(22725..22853)
    /rpt_family="MER5A"
  23951..24161
    /rpt_family="MER20"
  complement(24579..24655)
    /rpt_family="MIR"
  25020..25165
  
```

Query Match 5.8%; Score 57.8; DB 9; Length 99268;
 Best Local Similarity 67.9%; Pred. No. 0.0032;

```

Matches 112; Conservative 0; Mismatches 47; Indels 6; Gaps 2;
QY 829 AGGTGAGAGCGAAGAGCAGCAGCCATGTAGAAATGAACCGTCAGAGAGC--CAAGCAC 886
Db 65407 AGGTGAGAGCATACAAAGTAGCAGGTGTAGGATGAACAGGCTAGATCTACTGTATA 65466
QY 887 GGCAGAGGACTGCAGGCATCAGCGTGCATCTGTCGTTATTTGGAGTTCATGCCAAATGAG 946
Db 65467 ACATGAGACTGTAGGTAATAAATTGTACT---GTATGTGGAGTTCATGCTAAATGAG 65522
QY 947 TGTGTTTTAGTCTCTTGGCACAAAAAATAAAAAAAAAAAAAA 991
Db 65523 TAGATTTTAGTCTCTTGGCACAAACAAATAATCTAAGGGA 65567

RESULT 38
AL136374
LOCUS Human DNA sequence from clone RPI-244G5 on chromosome 1q24.3-25.3,
DEFINITION complete sequence.
ACCESSION AL136374
VERSION AL136374.4 GI:8919204
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 119853)
Kimberley,A.
Direct Submission
Submitted (17-JUL-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Jul 5, 2000 this sequence version replaced gi:8176662.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TrEMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr1
RPI-244G5 is from the library RPCI-1 constructed at the Roswell
Park Cancer Institute by the group of Pieter de Jong. For further
details see http://bacpac.med.buffalo.edu/
VECTOR: pCYPAC2
This sequence is the entire insert of clone RPI-244G5.
FEATURES
source
1..119853
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/map="q24.3-25.3"
/clone="RPI-244G5"
/clone_lib="RPCI-1"
repeat_region 1..237
/note="AluX repeat: matches 1..237 of consensus"
repeat_region 240..925
/note="L2 repeat: matches 1995..2706 of consensus"

```

```

complement(1507..1926)
/note="match: GSS: Em:AQ146487"
1822..2311
/note="match: GSS: Em:AQ624491"
1830..1956
/note="L2 repeat: matches 2616..2746 of consensus"
2127..2181
/note="L2 repeat: matches 2656..2710 of consensus"
2380..2691
/note="AluY repeat: matches 1..311 of consensus"
2716..3083
/note="MLT1B repeat: matches 1..390 of consensus"
complement(2776..3233)
/note="match: STS: Em:G37725
match: GSS: Em:B30422"
3252..3690
/note="match: GSS: Em:AQ773761"
4104..4288
/note="MIR repeat: matches 8..219 of consensus"
4533..4756
/note="L2 repeat: matches 1992..2228 of consensus"
4915..5144
/note="MIR repeat: matches 2..252 of consensus"
5284..5483
/note="L2 repeat: matches 2541..2749 of consensus"
7056..7103
/note="L2 copies 4 mer tcca 85% conserved"
7216..7391
/note="MIR repeat: matches 73..262 of consensus"
8076..8276
/note="MIR repeat: matches 61..262 of consensus"
8458..8494
/note="MIR repeat: matches 31..67 of consensus"
13489..13811
/note="L2 repeat: matches 1301..1648 of consensus"
13837..13892
/note="L2 copies 4 mer gtgt 75% conserved"
14392..15214
/note="MIR45B repeat: matches 1..835 of consensus"
16024..16224
/note="MIR repeat: matches 34..248 of consensus"
16428..16737
/note="AluY repeat: matches 1..299 of consensus"
complement(17098..17761)
/note="match: GSS: Em:AQ477207"
17714..18068
/note="L2 repeat: matches 2155..2537 of consensus"
18073..18710
/note="L2 repeat: matches 1049..1745 of consensus"
complement(18237..18836)
/note="match: GSS: Em:AQ527362"
complement(18332..19056)
/note="match: GSS: Em:AQ309729"
complement(18876..19056)
/note="match: STS: Em:? Em:HSPE98B10"
18880..19060
/note="match: STS: Em:? Em:HSPE23B05"
20092..20163
/note="L2 repeat: matches 2673..2746 of consensus"
20198..20725
/note="LTR8 repeat: matches 189..691 of consensus"
20726..20992
/note="AluSc repeat: matches 1..299 of consensus"
20993..21176
/note="LTR8 repeat: matches 1..189 of consensus"
21234..21886
/note="LTR1 repeat: matches 1..660 of consensus"
22078..22430
/note="LTR16C repeat: matches 3..385 of consensus"
22651..23063
/note="MIR1C repeat: matches 50..466 of consensus"
23449..23783
/note="MIR2 repeat: matches 1..345 of consensus"

```



```

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-498A13"
/clone_lib="RP11-11.2"
BASE COUNT 43651 a 26314 c 27269 g 41136 t
ORIGIN

```

```

Query Match          5.8%; Score 57.8; DB 9; Length 138370;
Best Local Similarity 56.8%; Pred. No. 0.0032;
Matches 107; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 800 GTTTGGGGGTTTCAGGATAGGGAATGGGAGGTCAGAGCAGCAGCAAGCAGCAGCCATGTA 859
Db 67409 GTTGGGGGTAGAGAGGAATGAGAAATTAGGTCAGAGGATGCAAGTAGTCATATATG 67468
QY 860 GAATGACCGTCAGAGAGCCAGCAGCGCAGAGGACTCCAGGCCATCAGCGTCGCACTGT 919
Db 67469 GGATGACACAGTCTAGAGATCTAATAAACAATAGGCTATAGTTACTAAAAATTGTA 67528
QY 920 TCGTATTTGGAGTTCATGCAGAAATGAGTGTGTTTAGCTGCTTCCACCAAAAAA 979
Db 67529 CTGTATTAGGATTATTGCTAATAAGTAGATTTTCAGCTGTTCTTCCCAAAAAACA 67588
QY 980 AAAAAA 988
Db 67589 CCAGAAAA 67597

```

```

RESULT 40
AC025375 156857 bp DNA linear HTG 25-MAY-2000
LOCUS Homo sapiens chromosome 1 clone RP11-2619 map 1, WORKING DRAFT
DEFINITION SEQUENCE, 15 unordered pieces.
ACCESSION AC025375
VERSION AC025375.3 GI:8077072
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 156857)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abrahams,H., Allen,N.,
Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collins,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrum,J., Meneus,P., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tefaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo.A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (08-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 25, 2000 this sequence version replaced gi:7328815.
All repeats were identified using RepeatMasker:

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abrahams,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bedalov,F.,
Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collins,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrum,J., Meneus,P., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tefaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo.A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (08-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 25, 2000 this sequence version replaced gi:7328815.
All repeats were identified using RepeatMasker:

```

```

Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L7958
Center clone name: 26119
----- Summary Statistics
Sequencing vector: M13, M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 149821 bases at least Q40
Consensus quality: 152963 bases at least Q30
Consensus quality: 154329 bases at least Q20
Insert size: 154000; agarose-fp
Insert size: 155457; sum-of-contigs
Quality coverage: 5.0 in Q20 bases; agarose-fp
Quality coverage: 5.0 in Q20 bases; sum-of-contigs
-----

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

```

1 778: contig of 778 bp in length
779 878: gap of 100 bp
879 1537: contig of 659 bp in length
1538 1637: gap of 100 bp
1638 4250: contig of 2613 bp in length
4251 4350: gap of 100 bp
4351 7895: contig of 3545 bp in length
7896 7995: gap of 100 bp
7996 11871: contig of 3876 bp in length
11872 11971: gap of 100 bp
11972 16256: contig of 4285 bp in length
16257 16356: gap of 100 bp
16357 21822: contig of 5466 bp in length
21823 21923: gap of 100 bp
21923 28321: contig of 6399 bp in length
28322 28421: gap of 100 bp
28422 36684: contig of 8263 bp in length
36685 36785: gap of 100 bp
36785 47215: contig of 10430 bp in length
47215 47314: gap of 100 bp
47314 59012: contig of 11698 bp in length
59013 59112: gap of 100 bp
59113 70244: contig of 11132 bp in length
70245 70344: gap of 100 bp
70345 86759: contig of 16415 bp in length
86760 86859: gap of 100 bp
86860 111698: contig of 24839 bp in length
111699 111799: gap of 100 bp
111799 156857: contig of 45059 bp in length.
FEATURES
source
1. 156857
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/map="1"
/clone="RP11-2619"
/clone_lib="RP11-11 Human Male BAC"
1. 778
/note="assembly_fragment
clone end:T7
vector side:right"
879. 1537
misc_feature
misc_feature

```

```

/note="assembly_fragment
clone end:SP6
vector side:left"
1638..4250
/note="assembly_fragment"
misc_feature
4351..7895
/note="assembly_fragment"
misc_feature
7996..11871
/note="assembly_fragment"
misc_feature
11972..16256
/note="assembly_fragment"
misc_feature
16357..21822
/note="assembly_fragment"
misc_feature
21923..28321
/note="assembly_fragment"
misc_feature
28422..36684
/note="assembly_fragment"
misc_feature
36785..47214
/note="assembly_fragment"
misc_feature
47315..59012
/note="assembly_fragment"
misc_feature
59113..70244
/note="assembly_fragment"
misc_feature
70345..86759
/note="assembly_fragment"
misc_feature
86860..111698
/note="assembly_fragment"
misc_feature
111799..156857
/note="assembly_fragment"
BASE COUNT 47296 a 31366 c 29768 g 47025 t 1402 others
ORIGIN
Query Match 5.8%; Score 57.8; DB 2; Length 156857;
Best Local Similarity 56.6%; Pred. No. 0.0032;
Matches 107; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
QY 800 GTTTGGGGGTTTCAGGATAGGAGTGGGGAGGTCAGAGAGCGCAAGCAGCAGCAGCATGTA 859
Db 52955 GTTGGGGGTAGAGAGGAATCCAGAAATTAGTTCAGAGGATGCGAAGTAGCATATATGTG 53014
QY 860 GAATGACCGTCCAGAGAGCCAGCAGGCGAGAGCTGCGAGGCCATCAGCGTCACGTGT 919
Db 53015 GGATGACCAAGCTAGAGATCTAATAACAACATAAGCGCTATAGTTACTAAAAATTGTA 53074
QY 920 TCGTATTGGAGTTTCATGCAGAAATCAGTGTCTTTTAGTGTCTTGGCACAAAAA 979
Db 53075 CTGTATTAGATTATGCTTAAATAAGTAGATTTCAGCTGTCTTGGCACAAAAA 53134
QY 980 AAAAAAAA 988
Db 53135 CCAGAAAAA 53143
RESULT 41
AC079289/c
LOCUS AC079289 161940 bp DNA linear HTG 20-AUG-2002
DEFINITION Homo sapiens chromosome 17 clone RP11-600J16 map 17, WORKING DRAFT
SEQUENCE, 7 unordered pieces.
ACCESSION AC079289
VERSION AC079289.2 GI:22325338
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Cararrhini; Hominidae; Homo.
1 (bases 1 to 161940)
Birren,B., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 17, clone RP11-600J16
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 161940)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,

```

Choepe, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heatford, A., Horton, L., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Labouque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lieu, G., Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., O'Neil, J., Peterson, K., Pierre, N., Pisan, C., Pollara, V., Raymond, C., Rieback, M., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Sougne, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission
Submitted (26-AUG-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 161940)
Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Chazaro, B., Choepe, Y., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gage, D., Gage, S., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (20-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 20, 2002 this sequence version replaced gi:9929721.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu

Center project name: L10440
Center clone name: 600 J 16
Project Information

Summary Statistics
Sequencing vector: M13; M77815; 27% of reads
Sequencing vector: Plasmid; n/a; 73% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 159446 bases at least Q40
Consensus quality: 160437 bases at least Q30
Consensus quality: 160863 bases at least Q20
Insert size: 157000; agarose-fp
Insert size: 161340; sum-of-contigs
Quality coverage: 11.9 in Q20 bases; agarose-fp
Quality coverage: 11.6 in Q2.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will

* be preserved.
 1 58877: contig of 58877 bp in length
 58878 58977: gap of 100 bp
 58978 60040: contig of 1063 bp in length
 60041 60140: gap of 100 bp
 60141 61238: contig of 1098 bp in length
 61239 61339: gap of 100 bp
 61339 65067: contig of 3729 bp in length
 65068 65168: gap of 100 bp
 65168 107152: contig of 41985 bp in length
 107153 107252: gap of 100 bp
 107253 154782: contig of 47530 bp in length
 154783 154883: gap of 100 bp
 154883 161940: contig of 7058 bp in length.

FEATURES

source
 1. .161940
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="17"
 /map="17"
 /clone="RP11-600J16"
 /clone_lib="RPC1-11 Human Male BAC"
 misc_feature
 1. .58877
 /notes="assembly_fragment"
 clone_end:SP6
 vector_side:left"
 misc_feature
 58978. .60040
 /notes="assembly_fragment"
 60141. .61238
 /notes="assembly_fragment"
 61339. .65067
 /notes="assembly_fragment"
 65168. .107152
 /notes="assembly_fragment"
 107253. .154782
 /notes="assembly_fragment"
 154883. .161940
 /notes="assembly_fragment"
 clone_end:T7
 vector_side:right"
 BASE COUNT 43128 a 36631 c 35390 g 46184 t 607 others
 ORIGIN

Query Match 5.8%; Score 57.8; DB 2; Length 161940;
 Best Local Similarity 67.9%; Pred. No. 0.0032;
 Matches 112; Conservative 0; Mismatches 47; Indels 6; Gaps 2;

QY 829 AGGTGAGGAGCGCAAGCAGCAGCGCATGTAGATGAACCGTCCAGAGAGC--CAAGCAC 886
 |||||
 Db 914 AGTTCAGAGCATACAAGTAGCAGGATGTAGGATGAACAAGCTAGAAATCTACTGTATA 855
 QY 887 GGCAGAGGACTGCGAGCGCCATCAGCGTGCACCTGTCTATTGAGTTTCATGCAAAATGAG 946
 |||||
 Db 854 ACATGAGGACTGTAGTATATAAATTGTACT---GTATGTGGAGTTCTATGCTAATGAG 799
 QY 947 TGTGTTTGTAGTCTCTTGGCCAAAAAATAAAAAAAAAAAAAA 991
 |||||
 Db 798 TAGATTGTAGTCTCTTGGCCAAACACAAAAAATCTAAGGAA 754
 |||||

RESULT 42
 AC107054/c
 LOCUS 50630 bp DNA linear PRI 29-MAY-2002
 DEFINITION Homo sapiens BAC clone RP11-400H10 from 4, complete sequence.
 AC107054
 ACCESSION
 AC107054.5 GI:20429604
 VERSION
 HTG.
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 50630)

AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 9847074
 REFERENCE
 2 (bases 1 to 50630)
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 3 (bases 1 to 50630)
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 4 (bases 1 to 50630)
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 5 (bases 1 to 50630)
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 6 (bases 1 to 50630)
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Sulston, J.E. and Waterston, R.
 Toward a complete human genome sequence
 Genome Res. 8 (11), 1097-1108 (1998)
 98053792
 2 (bases 1 to 50630)
 Wang, C., Kozlowicz, A., Spalding, L. and Boyer, E.
 The sequence of Homo sapiens BAC clone RP11-400H10
 Unpublished (2001)
 3 (bases 1 to 50630)
 Waterston, R.H.
 Direct Submission
 Submitted (14-JAN-2002) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 4 (bases 1 to 50630)
 Waterston, R.H.
 Direct Submission
 Submitted (04-APR-2002) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 5 (bases 1 to 50630)
 Waterston, R.H.
 Direct Submission
 Submitted (03-MAY-2002) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 6 (bases 1 to 50630)
 Waterston, R.
 Direct Submission
 Submitted (29-MAY-2002) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On May 3, 2002 this sequence version replaced gi:19924180.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: sapiens@wuston.wustl.edu
 ----- Summary Statistics

 Center project name: H_NH0400H10

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPC1-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Taten, M., Catanesi, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>
 VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-35D5, 2000 bp overlap; the

clone sequenced to the right is RP11-642E20, 2000 bp overlap.
Actual start of this clone is at base position 97704 of RP11-35D5.

FEATURES

```

source
1..50630
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP11-400H10"
/clone_lib="RC11-11"
151..630
/repeat_region
1007..1305
/repeat_region
2780..3061
/repeat_region
3241..3545
/repeat_region
4683..6142
/repeat_region
6143..6446
/repeat_region
6447..7296
/repeat_region
7297..7669
/repeat_region
7670..8151
/repeat_region
8411..8521
/repeat_region
8732..8767
/repeat_region
10181..10668
/repeat_region
10669..10936
/repeat_region
10937..11264
/repeat_region
12145..12226
/repeat_region
13486..13686
/repeat_region
14333..14356
/repeat_region
14395..14548
/repeat_region
14678..14849
/repeat_region
15282..15369
/repeat_region
15487..15510
/repeat_region
15511..15572
/repeat_region
15607..15753
/repeat_region
15982..16156
/repeat_region
16164..16390
/repeat_region
16406..16465
/repeat_region
16466..18051
/repeat_region
18052..18369
/repeat_region
18370..18939
/repeat_region
18960..19089
/repeat_region
19231..19323
/repeat_region

```

```

repeat_region 19322..20169
/rpt_family="L1"
repeat_region 20170..20479
/rpt_family="Alu"
repeat_region 20480..20630
/rpt_family="L1"
repeat_region 20800..21207
/rpt_family="MaLR"
repeat_region 21687..21731
/rpt_family="AT_rich"
repeat_region 21782..21840
/rpt_family="L2"
repeat_region 23361..23671
/rpt_family="Alu"
repeat_region 24503..24525
/rpt_family="AT_rich"
repeat_region 26923..27223
/rpt_family="Alu"
repeat_region 27248..27504
/rpt_family="Alu"
repeat_region 27930..27951
/rpt_family="(T)n"
repeat_region 28083..28152
/rpt_family="Mariner"
repeat_region 28172..28215
/rpt_family="AT_rich"
repeat_region 29024..29147
/rpt_family="L2"
repeat_region 29148..29264
/rpt_family="(TA)n"
repeat_region 29620..29682
/rpt_family="L2"
repeat_region 30174..30291
/rpt_family="L2"
repeat_region 30570..30623
/rpt_family="CRI"
repeat_region 30632..30692
/rpt_family="MER1_type"
repeat_region 30776..30821
/rpt_family="AT_rich"
repeat_region 31291..31420
/rpt_family="L1"
repeat_region 33262..33396
/rpt_family="MIR"

Query Match      5.8%; Score 57.4; DB 9; Length 50630;
Best Local Similarity 63.1%; Pred.No. 0.0041;
Matches 123; Conservative 0; Mismatches 66; Indels 6; Gaps 2;

QY 793 AGAGGAGTTTGGGGTTTCAGGATAGGGAATGGGAGGTTCAGAGGACGCAAGCAGCAG 852
Db 4996 AATGGTGGTTATAGGGGCAATTAATGGGAGATGTTCAGCCAGAGGATACAAAGTAGCAG 4937

QY 853 CCATGTAGATGAACCGTCCAGAGACCCTAA--GCACGGCAGGAGCTGAGGCGCATCAGC 910
Db 4936 ATATGTAGGATGGAAAGTCTGGAGATCTAATGTACAACTAGGACTATAGGTAATAAAA 4877

QY 911 GTCACTGTTTCGTATTGAGTTTCATGCAAAATGAGTGTGTTTAGCTGCTTGGCACA 970
Db 4876 TTGTA-----TTCTATGTAGGATTCATGCTAAATGATGATAGTATTTAGCTGCTGGCACA 4821

QY 971 AAAAAAAAAAAAAA 985
Db 4820 AAAAAACAAACAAAAA 4806

```

RESULT 43

```

AL357129
LOCUS
DEFINITION Human DNA sequence from clone RP11-114A21 on chromosome
Xq21.31-22.1 Contains ESTs, STSs and GSSs, complete sequence.
ACCESSION AL357129
VERSION AL357129.11 GI:10186549

```

KEYWORDS HTG. Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 145380)

AUTHORS Chapman,J.

TITLE Direct Submission

JOURNAL Submitted (19-DEC-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk

COMMENT requests: clonerequest@sanger.ac.uk

On Sep 19, 2000 this sequence version replaced gi:9909073.

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/ChrX>

This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known, annotated repeat sequence elements. Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key. RP11-114A21 is from the library RPCI-11.1 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACE3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-114A21 it may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.

The true left end of clone RP11-114A21 is at 1 in this sequence.

The true left end of clone RP11-122K23 is at 145281 in this sequence. The true left end of clone RP11-554A24 is at 80790 in this sequence. The true right end of clone RP11-361B11 is at 54224 in this sequence.

FEATURES

repeat_region	1..71_	Location/Qualifiers
	1..145380	
	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:9606"	
	/chromosome="X"	
	/map="q21.31-22.1"	
	/clone="RP11-114A21"	
	/clone_lib="RPCI-11.1"	
repeat_region	1..71_	
	/note="v1P3 repeat: matches 10. .80 of consensus"	
misc_feature	20..685	
	/note="match: GSS: Em:AQ343933"	
repeat_region	31..79	
	/note="v1P repeat: matches 4. .52 of consensus"	
repeat_region	450..432	
	/note="WER63 repeat: matches 2. .44 of consensus"	
repeat_region	847..993	
	/note="WER20 repeat: matches 29. .174 of consensus"	
repeat_region	1591..1614	
	/note="12 copies 2 mer tc 95% conserved"	
repeat_region	3441..3588	
	/note="2 copies 74 mer 92% conserved"	
repeat_region	3742..3801	
	/note="L2 repeat: matches 2643. .2701 of consensus"	
repeat_region	4195..4440	
	/note="WT1A1 repeat: matches 118. .365 of consensus"	
repeat_region	4442..4495	
	/note="HERV17 repeat: matches 1. .54 of consensus"	

(http://bacpac.med.buffalo.edu)

VECTOR: PRACE3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is AC021789; the clone sequenced to the right is AC021688. Actual start of this clone is at base position 1 of RP11-557N1; actual end is at base position 169202 of RP11-557N1.

The sequence H NH0557N01 contains a dinucleotide (GA) repeat from base position 29551 to 29868 where the sequence fidelity cannot be guaranteed. Assembly of the database is consistent with digest information.

FEATURES

source

Location/Qualifiers

```

1..169202
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-557N1"
/clone_lib="RPC1-11"
196..300
/rpt_family="Mariner"
407..730
/rpt_family="MER121"
891..934
/rpt_family="AT_rich"
1399..1428
/rpt_family="AT_rich"
1550..1577
/rpt_family="AT_rich"
1739..1895
/rpt_family="MIR"
3511..3635
/rpt_family="L2"
3700..4079
/rpt_family="ERV1"
4106..4201
/rpt_family="L2"
4947..5133
/rpt_family="MIR"
5960..5996
/rpt_family="polypurine"
6720..6772
/rpt_family="MIR"
6809..6889
/rpt_family="MIR"
7122..7348
/rpt_family="L1"
7349..7392
/rpt_family="GTTTG)n"
7393..7802
/rpt_family="L1"
7803..7856
/rpt_family="(TA)n"
7867..7984
/rpt_family="L1"
8738..8860
/rpt_family="MIR"
8863..9145
/rpt_family="Alu"
9426..9714
/rpt_family="Alu"
9715..9758
/rpt_family="AT_rich"
9903..9977
/rpt_family="L2"
10028..10056
/rpt_family="(TG)n"
10404..10689
/rpt_family="Alu"
12107..12278
/rpt_family="MaLR"

```

```

repeat_region 13854..13946
/rpt_family="ACHobo"
repeat_region 14020..14183
/rpt_family="MIR"
repeat_region 14531..14841
/rpt_family="Alu"
repeat_region 14860..15243
/rpt_family="MER2_type"
repeat_region 16032..16338
/rpt_family="Alu"
repeat_region 17440..17564
/rpt_family="CR1"
repeat_region 19909..19941
/rpt_family="AT_rich"
repeat_region 20077..20139
/rpt_family="MIR"
repeat_region 20386..20982
/rpt_family="ERV1"
repeat_region 21156..21524
/rpt_family="MaLR"
repeat_region 22038..22097
/rpt_family="MIR"
repeat_region 22230..22331
/rpt_family="MIR"
repeat_region 24135..24157
/rpt_family="AT_rich"
repeat_region 25901..25968
/rpt_family="MIR"
repeat_region 27874..28003
/rpt_family="MIR"
repeat_region 28292..28426
/rpt_family="MIR"
repeat_region 28611..28718
/rpt_family="L2"
repeat_region 28774..28835
/rpt_family="(TAAA)n"
repeat_region 29147..29261
/rpt_family="GA-rich"
repeat_region 29285..29567
/rpt_family="Alu"
repeat_region 29568..29642
/rpt_family="polypurine"
repeat_region 29644..29831
/rpt_family="GA-rich"
repeat_region 29832..29866
/rpt_family="(GGGA)n"
repeat_region 29991..30283
/rpt_family="Alu"
repeat_region 30686..31803
/rpt_family="L1"
repeat_region 31804..32105
/rpt_family="Alu"
repeat_region 32106..32750
/rpt_family="L1"

```

Query Match

Best Local Similarity 5.8%; Score 57.4; DB 9; Length 169202;

Matches 120; Conservative 0; Mismatches 61; Indels 6; Gaps 2;

```

Qy 805 GGGGTTTCAGGATAGGGAATGGGAGGTTCAGGAGCAGCAAGAGCAGCCATGTAGATG 864
Db 79581 GTGGGAGAGGAACTGGCGGGGTAGGTACAGGATAGCAATATGAGGATG 79640
Qy 865 AACGTCAGAGAGCC--AAGCAGGAGAGGATGAGGCCATCAGCGTGCATGTTCG 922
Db 79641 AACAGGCTGGAGACCTACAGCCCAACATACCCTACAAAGTATAAATCTGAGCT---G 79696
Qy 923 TATTGGAGTTTCATGCAAAATGACTGTGTTTACGCTCTTCCACAAAAAATAAAA 982
Db 79697 TATTGGAGTTTCATGCTGAATGGCTAGATTTCGCTCTTCCACAAAAAATAAAA 79756
Qy 983 AAAAAA 989

```

Db 79757 TGAATAA 79763

RESULT 45
AC145171/c

LOCUS
DEFINITION
AC145171 Homo sapiens chromosome 16 clone RP11-349F16, WORKING DRAFT
SEQUENCE, 3 unordered pieces.

ACCESSION
AC145171 GI:31621312

VERSION
HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.

KEYWORDS
Homo sapiens (human)

SOURCE
Homo sapiens

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 175947)
DOE Joint Genome Institute.
Sequencing of Human Chromosome 16
Unpublished
2 (bases 1 to 175947)
DOE Joint Genome Institute.
Direct Submission
Submitted (12-JUN-2003) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

COMMENT
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: <http://www.jgi.doe.gov>

Project Information
Center Project Name: 540517
Center clone name: RP11-11_349F16

Summary Statistics
Consensus quality: 173647 bases at least Q40
Consensus quality: 174869 bases at least Q30
Consensus quality: 175338 bases at least Q20
Estimated insert size: 175000; agarose-fp estimation
Quality coverage: 9.21 in Q20 bases; agarose-fp estimation
Quality coverage: 9.17 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
* 1 45369: contig of 45369 bp in length
* 45370 45469: gap of unknown length
* 45470 94891: contig of 49422 bp in length
* 94892 94991: gap of unknown length
* 94992 175947: contig of 80956 bp in length.

FEATURES
source
Location/Qualifiers
1..175947
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-349F16"
/clone_lib="RPC1 human BAC library 11"
BASE COUNT 54782 a 31550 c 31265 g 58150 t 200 others
ORIGIN

Query Match 5.8%; Score 57.4; DB 2; Length 175947;
Best Local Similarity 55.9%; Pred. No. 0.004;
Matches 109; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 795 GAGGAGTTTGGGGGTTTCAGGATAGGGAATGGGAGGTCAGAGGACGCAAGCAGCACC 854
Db 39798 GTGGGATGGAGTGGGACTAATAGGAGGAGGTTAGGTAGGCTTGCACAAAGTAGCAT 39739

QY 855 ATGTAGTAATGACCGTCCAGAGAGCCAGCGGAGAGGACTGCAGGCCATCAGCGTGC 914
Db 39738 AAGTAGGATGACAAAGTCTAGAGATCTTAATGTACAATATGAGAACTATAGCTTTAAAAAA 39679

QY 915 ACTGTTCTGATTTGGAGTTCATGCAAAATGAGTGTGTTTTAGTCTCTTTCGCCACAAAA 974
Db 39678 ATTGTATTGTATTGGTTTCTGTGTTAAATGAGTAGCTTTAGTCTCTTTCGTACAAAT 39619

QY 975 AAAAAAAGAAAAA 989
Db 39618 CAACAACAAGAAAA 39604

AC013408 151841 bp DNA linear PRI 09-JAN-2002
Homo sapiens BAC clone RP11-469G18 from 2, complete sequence.

AC013408
AC013408
AC013408.7 GI:15920142
HTG.
Homo sapiens (human)

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 151841)
Sulston, J.E. and Waterston, R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
99063792
9847074

REFERENCE
2 (bases 1 to 151841)
Holmes, A., Hawkins, M. and Kozlowicz, A.
The sequence of Homo sapiens BAC clone RP11-469G18
Unpublished (2001)
3 (bases 1 to 151841)
Waterston, R.H.
Direct Submission
Submitted (09-NOV-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 151841)
Waterston, R.H.
Direct Submission
Submitted (04-OCT-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 151841)
Waterston, R.
Direct Submission
Submitted (09-JAN-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Oct 4, 2001 this sequence version replaced gi:13992768.

COMMENT
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@wustl.wustl.edu
----- Summary Statistics
----- Center project name: H_NH0469G18

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tateno, M., Catanese, J.J., and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-90C9, 2000 bp overlap; the clone sequenced to the right is RP11-22L19, 2000 bp overlap. Actual start of this clone is at base position 1 of RP11-469G18; actual end is at base position 14496 of RP11-22L19.

Data from AC079146 was used to finish this clone, AC013408.

Polymorphisms have been identified between AC079146 and AC013408.

FEATURES

source

Location/Qualifiers

1. .151841

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosomes="2"

/map="2"

/clone="RP11-469G18"

/clone_lib="RPCI-11"

2. .1209

/rpt_family="ERVVK"

1210. .2222

/rpt_family="ERVVK"

3426. .3844

/rpt_family="L1"

3559. .3593

/rpt_family="CA)n"

5300. .5476

/rpt_family="MIR"

5483. .5601

/rpt_family="ACHobo"

6149. .6256

/rpt_family="L2"

6526. .6700

/rpt_family="MaLR"

9037. .9184

/rpt_family="(TA)n"

9834. .9864

/rpt_family="(GA)n"

12527. .12597

/rpt_family="(TA)n"

13739. .13760

/rpt_family="AT_rich"

13964. .13999

/rpt_family="AT_rich"

14037. .14242

/rpt_family="MaLR"

14214. .14240

/rpt_family="(TC)n"

14483. .14787

/rpt_family="Alu"

15016. .15042

/rpt_family="(TAGA)n"

15269. .15555

/rpt_family="Alu"

17128. .17147

/rpt_family="(TTTTTA)n"

17131. .17430

```
repeat_region      /rpt_family="Alu"
17445. .17472
repeat_region      /rpt_family="(T)n"
17760. .17782
repeat_region      /rpt_family="AT_rich"
17859. .17931
repeat_region      /rpt_family="MIR"
18428. .18484
repeat_region      /rpt_family="MaLR"
18474. .18496
repeat_region      /rpt_family="AT_rich"
18485. .19089
repeat_region      /rpt_family="L1"
19090. .19431
repeat_region      /rpt_family="MaLR"
19465. .19547
repeat_region      /rpt_family="GA_rich"
22482. .22512
repeat_region      /rpt_family="AT_rich"
22637. .22734
repeat_region      /rpt_family="L1"
29040. .29078
repeat_region      /rpt_family="(CA)n"
29467. .29576
repeat_region      /rpt_family="MIR"
29686. .29874
repeat_region      /rpt_family="MIR"
30573. .30990
repeat_region      /rpt_family="ERVL"
30605. .30627
repeat_region      /rpt_family="AT_rich"
31817. .31898
repeat_region      /rpt_family="(TA)n"
31979. .32006
repeat_region      /rpt_family="AT_rich"
32107. .32423
repeat_region      /rpt_family="Alu"
32397. .32427
repeat_region      /rpt_family="(TAAA)n"
33618. .33656
repeat_region      /rpt_family="AT_rich"
35713. .36026
repeat_region      /rpt_family="MER1_type"
35814. .35845
repeat_region      /rpt_family="AT_rich"
35904. .35947
repeat_region      /rpt_family="AT_rich"
37385. .37417
repeat_region      /rpt_family="AT_rich"
37594. .38042
repeat_region      /rpt_family="L1"
37733. .37788
repeat_region      /rpt_family="(TG)n"
38010. .38313
repeat_region      /rpt_family="L1"
38406. .38429
repeat_region      /rpt_family="(TTG)n"
38847. .38946
repeat_region      /rpt_family="MIR"
38947. .39247
repeat_region      /rpt_family="Alu"
39248. .39385
repeat_region      /rpt_family="MIR"
39419. .39650
repeat_region      /rpt_family="L1"
41061. .41384
repeat_region      /rpt_family="L1"
41414. .41583
repeat_region      /rpt_family="MIR"
```

Query Match 5.7%; Score 56.6; DB 9; Length 151841;
 Best Local Similarity 56.0%; Pred. No. 0.0062;
 Matches 107; Conservative 0; Mismatches 84; Indels 0; Gaps 0;


```

repeat_region complement(4673..4745)
repeat_region /rpt family="MIR"
repeat_region complement(4765..4938)
repeat_region /rpt family="MIR"
repeat_region 6220..6268
repeat_region /rpt family="AT_rich"
repeat_region complement(6368..6543)
repeat_region /rpt family="MIR"
repeat_region 7369..7664
repeat_region /rpt family="AluSx"
repeat_region complement(7959..8673)
repeat_region /rpt family="LTR1B"
repeat_region 9636..9686
repeat_region /rpt family="MIR"
repeat_region 10618..10638
repeat_region /rpt family="AT_rich"
repeat_region complement(11204..11513)
repeat_region /rpt family="AluSg"
repeat_region 11624..12119
repeat_region /rpt family="MLT1D"
repeat_region complement(12125..12194)
repeat_region /rpt family="MIR"
repeat_region 12717..12757
repeat_region /rpt family="AT_rich"
repeat_region complement(13082..13389)
repeat_region /rpt family="AluY"
repeat_region 15244..15324
repeat_region /rpt family="AT_rich"
repeat_region complement(15341..15853)
repeat_region /rpt family="L1M2"
repeat_region 15861..15969
repeat_region /rpt family="FRAM"
repeat_region 15970..15992
repeat_region /rpt family="(TTG)n"
repeat_region complement(15997..16293)
repeat_region /rpt family="AluSp"
repeat_region complement(16318..16379)
repeat_region /rpt family="L1M4"
repeat_region 16417..16534
repeat_region /rpt family="FLAM_C"
repeat_region 16672..16697
repeat_region /rpt family="AT_rich"
repeat_region complement(16770..16958)
repeat_region /rpt family="L1M4"
repeat_region complement(16962..17095)
repeat_region /rpt family="L1M2"
repeat_region complement(17096..17477)
repeat_region /rpt family="L1M4"
repeat_region 17514..17534
repeat_region /rpt family="AT_rich"
repeat_region complement(17535..17879)
repeat_region /rpt family="L1M4a"
repeat_region 17880..17910
repeat_region /rpt family="(TA)n"
repeat_region complement(17911..18043)
repeat_region /rpt family="L1M4a"
repeat_region complement(18667..18983)
repeat_region /rpt family="L1M2"
repeat_region 19055..19102
repeat_region /rpt family="(TG)n"
repeat_region 19398..19494
repeat_region /rpt family="T-rich"
repeat_region 20011..20043
repeat_region /rpt family="AT_rich"
repeat_region complement(20337..20507)
repeat_region /rpt family="HAL1b"
repeat_region note="c30 qual SNGL region"
repeat_region 20966..20995
repeat_region /rpt family="(TG)n"
repeat_region 20995..21020
repeat_region /rpt family="(GA)n"
repeat_region 22075..22119

```

```

/rpt family="AT_rich"
Query Match 5.7%; Score 56.6; DB 9; Length 173270;
Best Local Similarity 66.9%; Pred. No. 0.0062;
Matches 113; Conservative 0; Mismatches 49; Indels 7; Gaps 2;

QY 814 GGATAGGAGATGGGAGGTGAGAGCGCAAGAGCAGCCATGTAGAAATGAAC---CGT 870
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 871 CCAGAGAGCCAGCAGCGCAGAGCACTGCAGGCCATCAGCGTGCACGTTCGTATTGGA 930
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 42134 AGAGATCTAATATACGATATGAGGATATAGGTAATAACATTCACAT---GTATTGGA 42189
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 931 GTTCATGCAAAATGAGTGTGTTTGTAGCTGCTTGCACAAAAA 979
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 42190 ATTCAATGCTAAATGAGTAGATTTTGTAGCTGTTCTTGCACAAAAA 42238
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 48
AC016256 164405 bp DNA linear PRI 25-JUL-2001
LOCUS Homo sapiens 12 BAC RP11-98B6 (Roswell Park Cancer Institute Human
DEFINITION BAC library) complete sequence.
ACCESSION AC016256
VERSION AC016256.19 GI:14670071
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 164405)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alzbrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbara,J.,
Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowle,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D.,
Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.I., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D.,
Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J.,
Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W.,
Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K.,
Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O.,
Hodgson,A., Hogue,M., Holloway,C., Hollins,B., Honsi,F.,
Howard,S., Huber,J., Hulyk,S., Hume,J., Ioshikhes,I., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulis,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Marondel,I., Martin,R.,
Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P.,
Meador,M., Mei,G., Merscher,S., Metzger,M., Miller,A., Miner,G.,
Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M.,
Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
Ogih,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojokan,I., Roife,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shimm,C.,
Shoostari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A.,
Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A.,
Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B.,
Thomas,N., Thomas,S., Usmami,K., Vasquez,L., Vera,V., Villalon,D.,
Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R.,
Washington,C., Watlington,S., Williams,G., Williamson,A.,
Wlecyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J.,
Zorrilla,S., Kucherlapati,R. and Gibbs,R.

```



```

/function="Polymorphic site"
14426

misc_feature
Query Match
Best Local Similarity 5.7%; Score 56.4; DB 9; Length 164405;
Matches 137; Conservative 0; Mismatches 96; Indels 3; Gaps 2;

QY 757 CTTGCCCGCTACAGGAGCACCGCCGCTCTGAGTGAAGAGGAGTTGTGGGGGTTCCAGGA 816
| | | | |
Db 143164 CTTGAATATATAGAGATAGACGAAGAACGACGTGGATACAGCTGGAGAGTGGTCCAAA 143223

QY 817 TAGGGAATGGGAGGTCAGAGACCGCAAGCAGCAGCCCATGTAGATGAACCGTCCAGAG 876
| | | | |
Db 143224 ATGGGAAGATGTAGTATTAGAGGATACAAATTA--AGTTATGTAGGATGAACAAGTCTAGA 143281

QY 877 AGCCAAGCAGCGCAGAGGACTGCAGGCCATCAGCGTGCACCTGTCGTATTGGAGTTTCAT 936
| | | | |
Db 143282 GGTCTAAGATACATATCAGGACATAAATATATAAAATTGTACT-GTGTCTGTGATTTCAT 143340

QY 937 GCAAAATGAGTGTGTTTGTAGTCTGCTTGGCCACAAAAAAGAAAAAAGAAAAA 992
| | | | |
Db 143341 GCTAAATGAGTCGATTTTGTAGTCTGCTTGGCCACAAACCAAAAAAAGAAAAA 143396

RESULT 49
AC025258/c
LOCUS
DEFINITION
Homo sapiens chromosome 12 clone RP11-398017, WORKING DRAFT
SEQUENCE, 2 unordered pieces.
AC025258
AC025258.15 GI:14547352
HTG; HTGS PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 16668)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Alsbrooke,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbara,J.,
Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Bunay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C.,
Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulsegh,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Nowatson,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokkwo,S.,
Ogih,M., Okuwonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshkari,N.,
Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.

```

```

Db      111101 GCTAATGAGTCGATTTAGTCTCTTGCACAAACCAAAAAAAAAAAAAAAAA 111046

RESULT 50
LOCUS   AC011004/c
DEFINITION Homo sapiens chromosome 8 clone RP11-110D15, WORKING DRAFT
ACCESSION AC011004
VERSION   AC011004.12 GI:13273359
KEYWORDS  HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FUILLTOP.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens

REFERENCE
AUTHORS   1 (bases 1 to 183778)
           Abola, A.P., Bruno, D., Conn, L., Dela Rosa, M., Faulkner, D.,
           Federspiel, N., Glukhov, S., Hansen, N., Herman, Z.S., Hyman, R.,
           Mao, J., Komp, C., Kottler, S., Lam, B., Marathe, R., Miranda, M.,
           Morehouse, A.J., Nguyen, M., Oefner, P., Palm, C.J., Ramirez, D.,
           Southwick, A.M., Webb, C., Wilhelm, J., Yu, S. and Davis, R.W.
           Unpublished
JOURNAL   2 (bases 1 to 183778)
           Bruno, D., Conn, L., Dela Rosa, M., Federspiel, N., Foreman, P.,
           Glukhov, S., Hansen, N., Hyman, R., Mao, J., Marathe, R.,
           Morehouse, A.J., Oefner, P., Palm, C.J., Ramirez, D. and Davis, R.W.
           Direct Submission
TITLE     Submitted (29-SEP-1999) DNA Sequencing and Technology Center,
JOURNAL   Stanford University, 855 California Avenue, Palo Alto, CA 94304,
           USA

COMMENT   On Mar 11, 2001 this sequence version replaced gi:13270571.
           ----- Genome Center
           Center: Stanford DNA Sequencing and Technology Development
           Center
           Center code: SOSTDC
           Web site: http://sequence-www.stanford.edu/group/human/
           Contact: hum-info@sequence.stanford.edu
           ----- Project Information
           Center project name: 666
           Center clone name: RP11-110D15
           ----- Summary Statistics
           Sequencing Vector: M13mp18; X02513
           Chemistry: Dye-primer; 0% of reads
           Assembly: Dye-terminator Big Dye; 100% of reads
           Assembly program: Phrap; version 0.990319
           Consensus quality: 182108 bases at least Q40
           Consensus quality: 183015 bases at least Q30
           Consensus quality: 183336 bases at least Q20
           Insert size: 185712; agarose-fp
           Quality coverage: 9.3x in Q20 bases; agarose-fp
           Quality coverage: 9.4x in Q20 bases; sum-of-contigs.
           * NOTE: This is a 'working draft' sequence. It currently
           * consists of 3 contigs. The true order of the pieces
           * is not known and their order in this sequence record is
           * arbitrary. Gaps between the contigs are represented as
           * runs of N, but the exact sizes of the gaps are unknown.
           * This record will be updated with the finished sequence
           * as soon as it is available and the accession number will
           * be preserved.
           *
           * 1 26915: contig of 26915 bp in length
           * 26916 27015: gap of unknown length
           * 27016 98873: contig of 71858 bp in length
           * 98874 98973: gap of unknown length
           * 98974 183778: contig of 84805 bp in length.
           Location/Qualifiers
           1. 183778
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"
              /chromosome="8"
              /clone="RP11-110D15"
              /clone_lib="RPCI human BAC library 11"

```

```

misc_feature 1..26915
              /note="assembly_name:Contig8
              clone_end:SP6"
              27016..98873
              /note="assembly_name:Contig9
              clone_end:T7"
              98974..183778
              /note="assembly_name:Contig10"
              201 others
BASE COUNT 50361 a 41935 c 41401 g 49880 t
ORIGIN

Query Match      5.6%; Score 56; DB 2; Length 183778;
Best Local Similarity 64.4%; Pred. No. 0.0087;
Matches 116; Conservative 0; Mismatches 60; Indels 4; Gaps 2;

QV      813 AGGATAGGGAATGGGGAGGTGAGAGGAGCGCAAAAGCAGCAGCATGTAGAAATGAACCGTCC 872
Db      37638 AGGAAACGGGAGATCTAGTGTGAAGGACGACAGAGGAGGAGATATGCAAGATGCATCGCCC 37579
QV      873 AGAGAGCCCAAGCAGCGGACAGAGGACTGCGAGGCGCATCGCGTGCACCTGTTTCGTATTGGAGT 932
Db      37578 AGAGACCTTAATGT-GCATGAGGAGCTCTCTAGTCAACCAACATTGA---ATTATACCGGGGAT 37523
QV      933 TCATGCAAAATGAGTGTGTTTGTAGTCTGCTCTTGCACAAACCAAAAAAAAAAAAAAAAAAAAA 992
Db      37522 TTCTGCTAAGACAGTAGATTTTAGGAGCTCTTGGCATATTGCAAAAAAAAAAAAAAAAAAAAA 37463

Search completed: February 9, 2004, 14:16:33
Job time : 3928 secs

```

FEATURES
source

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2004, 13:03:52 ; Search time 2524 Seconds
(without alignments)
9552.314 Million cell updates/sec

Title: US-09-990-726-222

Perfect score: 992

Sequence: 1 ggcacgacgacgaactagg.....aaaaaaaaaaaaaaaaaaaa 992

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

EST:*

1: em_estba:*

2: em_estbum:*

3: em_estin:*

4: em_estnu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_hic:*

9: gb_esti:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vri:*

28: gb_gssl:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	626.8	63.2	694	13	BX097129 BX097129
C 2	409.8	41.3	414	14	H74303 YU56C11.S1
C 3	408	41.1	437	14	N54458 YV40B02.S1
4	373.2	37.6	480	14	H73373 YU48F10.R1

5	354	35.7	418	14	H58326	H58326 Yr25c07.r1
6	349	35.2	582	14	H74302	H74302 YU56C11.R1
7	340.6	34.3	468	14	N76568	N76568 YV40B02.R1
8	320.4	32.3	996	11	AK020081	AK020081 Mus muscu
C 9	286	28.8	405	9	AI242058	AI242058 qh81d07.x
C 10	284.2	28.6	296	14	H73374	H73374 YU48F10.S1
C 11	277.2	27.9	379	14	R02548	R02548 ye80a07.s1
C 12	274.2	27.6	450	9	AI438986	AI438986 t884a08.x
C 13	268.8	27.1	553	14	BY718164	BY718164 RZPD Mus
C 14	265.2	26.7	528	4	BX529784	BX529784 RZPD Mus
C 15	261.2	26.3	384	14	H40263	H40263 YP59E11.S1
C 16	256.8	25.9	284	14	H58715	H58715 Yr25c07.S1
C 17	235.2	23.7	376	14	T80187	T80187 YD85G03.S1
C 18	227.4	22.9	334	9	AA676653	AA676653 ZJ67E03.S
C 19	227	22.2	358	9	AI022453	AI022453 OW96H04.X
C 20	220.2	22.2	358	14	R02666	R02666 ye80a07.r1
C 21	218.4	22.0	1161	14	W41959	W41959 mc68b10.r1
C 22	202	20.4	401	9	AA050499	AA050499 m320F02.F
C 23	192.2	19.4	324	14	T87491	T87491 YD85G03.R1
C 24	182.2	18.4	467	14	W65713	W65713 me09G06.R1
C 25	175.6	17.7	218	14	N86939	N86939 L1498F Huma
C 26	174.6	17.6	364	13	BY212485	BY212485 BY212485
C 27	154	15.5	537	12	BI347002	BI347002 376394 MA
C 28	154	15.5	582	12	BI347003	BI347003 376395 MA
C 29	153.8	15.5	443	14	CB051335	CB051335 NISC_GJ23
C 30	153.8	15.5	540	14	CB051334	CB051334 NISC_GJ23
C 31	145.2	14.6	194	14	N88474	N88474 K3475F Huma
C 32	143	14.4	382	10	BB842514	BB842514 BB842514
C 33	132.6	13.4	370	10	BB843270	BB843270 BB843270
C 34	93.2	9.4	420	10	BF443482	BF443482 260980 MA
C 35	92.6	9.3	503	10	BF442092	BF442092 258654 MA
C 36	79.8	8.0	208	9	AA031034	AA031034 m147H05.F
C 37	57.8	5.8	483	9	AI187208	AI187208 gf28B11.X
C 38	52.8	5.3	466	9	AW794912	AW794912 RC6-UM001
C 39	52.6	5.3	493	2	HSNC77193	Bx487006 Homo sapi
C 40	51.2	5.2	606	12	BI063151	BI063151 IL3-UT011
C 41	50.8	5.1	443	13	BY560588	BY560588 BY560588
C 42	49	4.9	925	29	CNS00912P	AL053013 Drosophila
C 43	48	4.8	844	29	CNS00521P	AL056652 Drosophila
C 44	47.4	4.8	974	13	AW795001	AW795001 RC6-UM001
C 45	47.2	4.8	974	13	BX339244	BX339244 BX339244
C 46	46.4	4.7	748	29	BZ772737	BZ772737 mcv51f02
C 47	46.2	4.7	1201	13	BX356664	BX356664 BX356664
C 48	45.6	4.6	569	14	CB615918	CB615918 AMGNNUC:U
C 49	45.6	4.6	626	14	CD397038	CD397038 Gm_CK1771
C 50	45.6	4.6	935	29	CNS006XK	AL066051 Drosophila
C 51	45.4	4.6	925	29	CNS00912P	AL053013 Drosophila
C 52	45.4	4.6	1201	13	BX381961	BX381961 BX381961
C 53	45	4.5	537	14	CA442247	CA442247 UI-H-D10-
C 54	44.8	4.5	739	29	AG187401	AG187401 Pan trogl
C 55	44.6	4.5	638	14	CB242205	CB242205 UI-CF-PN0
C 56	44.2	4.5	1100	29	CNS00FSE	AL070988 Drosophila
C 57	43.4	4.4	660	13	BU928339	BU928339 AGENCOURT
C 58	43.2	4.4	420	9	AA1333438	AA1333438 S21G11 AG
C 59	43	4.3	449	9	AA186675	AA186675 ZP71C08.S
C 60	43	4.3	591	14	CB000366	CB000366 S345U_G08
C 61	43	4.3	699	29	AG169292	AG169292 Pan trogl
C 62	43	4.3	940	13	BUI185569	BUI185569 AGENCOURT
C 63	42.6	4.3	349	28	AQ097268	AQ097268 HS 3038 B
C 64	42.6	4.3	524	28	AQ480601	AQ480601 RPCI-11-B
C 65	42.6	4.3	628	29	AG018533	AG018533 Homo sapi
C 66	42.6	4.3	638	29	AG018534	AG018534 Homo sapi
C 67	42.6	4.3	639	29	AG017811	AG017811 Homo sapi
C 68	42.6	4.3	641	29	AG017810	AG017810 Homo sapi
C 69	42.6	4.3	643	29	AG017808	AG017808 Homo sapi
C 70	42.4	4.3	699	14	CB046795	CB046795 NISC_GF06
C 71	42.2	4.3	385	9	AI478457	AI478457 tm44C09.X
C 72	42.2	4.3	469	13	BQ169652	BQ169652 WHE0961_B
C 73	42.2	4.3	1036	9	AL550172	AL550172 AL550172
C 74	42	4.2	543	9	AV597018	AV597018 AV597018
C 75	42	4.2	549	28	AQ221080	AQ221080 HS 3223 B
C 76	41.8	4.2	260	10	BF763246	BF763246 RC6-CS007
C 77	41.8	4.2	289	9	AW665077	AW665077 hi86h11.X


```

Db      125  ACTGCTACCTACTTCTGCGGCGCTCCACCTCAGGTGCCCATGTGGACAGTGCCA 184
QY      403  GGCTACAGATGCACTGGGAGCTGTGTGTCACAGCCAGTGTCTGAGCTGCGGGCCAACTTCA 462
Db      185  GGCTACAGATGCACTGGGAGCTGTGTGTCACAGCCAGTGTCTGAGCTGCGGGCCAACTTCA 214
QY      463  CTCTGACGACAGAGGGGCGAGCCCGCCAGGCTGGAGATGATCTGCCAGGGGCTCTCGGGCA 522
Db      215  -----GACAGAGGGGCGAGCCCGCCAGGCTGGAGATGATCTGCCAGGGGCTCTCGGGCA 266
QY      523  GCCACACTATCACCAACAGCCTCATCGGGAAGATGGCGAGTCCACCTGCACAGAGAC 582
Db      267  GCCACACTATCACCAACAGCCTCATCGGGAAGATGGCGAGTCCACCTGCACAGAGAC 326
QY      583  CATGCCACAGGCGCTGCCAACTTCTCTCTGCGGAGCCAGACATCGGACTGTCTTCT 642
Db      327  CATGCCACAGGCGCTGCCAACTTCTCTCTGCGGAGCCAGACATCGGACTGTCTTCT 386
QY      643  GGTGCCAGGCTGCAACCAAGCCCAATGTCCAGCACAGCCCTTACAGTGGTGCCTCCAG 702
Db      387  GGTGCCAGGCTGCAACCAAGCCCAATGTCCAGCACAGCCCTTACAGTGGTGCCTCCAG 446
QY      703  GTGGTGACCAAGATGGAGGACTGCGAGGTCCTCCCTGGAGAGCCCTTCTGCTTGC 762
Db      447  GTGGTGACCAAGATGGAGGACTGCGAGGTCCTCCCTGGAGAGCCCTTCTGCTTGC 506
QY      763  CGCTCTACAGGAGCACCCGCCCTCTGAGTGAAGAGGATTTGGGGGTTTCAGATAGGCA 822
Db      507  CGCTCTACAGGAGCACCCGCCCTCTGAGTGAAGAGGATTTGGGGGTTTCAGATAGGCA 566
QY      823  AT-GGGAGCTCAGAGGACGCAAGACAGCAGCCATGTAGATGAACCGTCCAGAGAGCA 881
Db      567  ATGGGGGAGGTGAGAGGACGCAAGACAGCAGCCATGTAGATGAACCGTCCAGAGAGCA 626
QY      882  AGCAGGCGACAGACTGCGAGGCGCATCAGGTGCTGCTGTTGCTATTGGAGTTTCATGCAA 941
Db      627  AGCAGGCGACAGACTGCGAGGCGCATCAGGTGCTGCTGTTGCTATTGGAGTTTCATGCAA 686
QY      942  ATGAGTGT 949
Db      687  ATGAGTGT 694

```

```

RESULT 2
H74303/c
LOCUS
DEFINITION
IMAGE:230132 3', mRNA sequence.
ACCESSION
H74303.1 GI:1047714
VERSION
H74303.1
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 414)
Hillier, L., Lemon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,
Chasse, S., Dietrich, N., DuBuque, T., Favell, A., Gish, W., Hawkins,
M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore,
B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T.,
Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E.,
Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
JOURNAL
MEDLINE
PUBMED
889549
CONTACT: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

```

```

Insert Size: 1114
High quality sequence stops: 313
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1114 Std Error: 0.00
Seq primer: Promega -21ml3
High quality sequence stop: 313.
Location/Qualifiers
1. .414
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GD3:3781228"
/db_xref="taxon:9606"
/clone="IMAGE:230132"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen INFLS"
/note="Organ: Liver and Spleen; Vector: p7T3D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
15' AACTGGAAGAATAAATAGATCTTTTTTTTTTTT 3',
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified p7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 75 a 127 c 116 g 95 t 1 others
ORIGIN
Query Match 41.3%; Score 409.8; DB 14; Length 414;
Best Local Similarity 99.3%; Pred No. 2.6e-48;
Matches 411; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 556 ATGGCAGGTCACCTGCGAGCAGACCATCCACAGGAGCCTTCTCTTCC 615
Db 414 ATGGCAGGTCACCTGCGAGCAGACCATCCACAGGAGCCTTCTCTTCC 355
QY 616 TGCCGAGCCAGACATCGGACTGTTCTGTCAGGCTGCAACACGCGCATGTCAGC 675
Db 354 TGCCGAGCCAGACATCGGACTGTTCTGTCAGGCTGCAACACGCGCATGTCAGC 295
QY 676 ACAGCGCCCTCACAGTGTGTCGCCCGCAGGTGTGTACCAAGATGAGGACTGGCAGGTC 735
Db 294 ACAGCGCCCTCACAGTGTGTCGCCCGCAGGTGTGTACCAAGATGAGGACTGGCAGGTC 235
QY 736 CCCTGGAGAGCCCATCTTTCCTTGGCGCTCTACAGGAGCACCGCGCTCTGAGTGAAG 795
Db 234 CCCTGGAGAGCCCATCTTTCCTTGGCGCTCTACAGGAGCACCGCGCTCTGAGTGAAG 175
QY 796 AGGAGTTGGGGGCTCAGGATAGGATGGGAGTTCAGGAGCGCAACAGCAGCAGCA 855
Db 174 AGGAGTTGGGGGCTCAGGATAGGATGGGAGTTCAGGAGCGCAACAGCAGCAGCA 115
QY 856 TGTAAGATGAACCTCCAGAGAGCCAGACCGCAGAGGACTGCGAGCCATCAGCGTCA 915
Db 114 TGTAAGATGAACCTCCAGAGAGCCAGACCGCAGAGGACTGCGAGCCATCAGCGTCA 55
QY 916 CTGTTCTGATTGAGTTCATGCAAAATGAGTGTGTTTTAGCTGCTCTTGCCAC 969
Db 54 CTGTTCTGATTGAGTTCATGCAAAATGAGTGTGTTTTAGCTGCTCTTGCCAC 1

```

```

RESULT 3
N54458/c
LOCUS
DEFINITION
IMAGE:245163 3', mRNA sequence.
ACCESSION
N54458
VERSION
N54458.1 GI:1195778
KEYWORDS
EST.
SOURCE
Homo sapiens (human)

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 437)
REFERENCE
AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,
Chisoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins,
M., Hultman, M., Kucaba, T., Lacy, M., Le, N., Le, N., Mardis, E., Moore,
B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T.,
Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E.,
Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
97044478
PUBMED
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLMNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 832 Std Error: 0.00
Seq primer: ml3 -40 forward
High quality sequence stop: 375.
Location/Qualifiers
1..437
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3794409"
/db_xref="taxon:9606"
/clone="IMAGE:245163"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen INFLS"
/note="Organ: Liver and Spleen; Vector: p773D (Pharmacia)
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5' AACTGGAAGTAATTAAGATCTTTTATTTTATTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 78 a 132 c 119 g 106 t 2 others
ORIGIN
Query Match 41.1%; Score 408; DB 14; Length 437;
Best Local Similarity 97.3%; Pred No. 4.5e-48;
Matches 425; Conservative 0; Mismatches 11; Indels 1; Gaps 1;
QY 537 AACAGCCTGATCGGAAGATGGGACAGTCCACCTGCAGCAGAGACCATGCCACAGGCAG 596
Db 437 AACAGCCTGATCGGAAGATGGGACAGTCCACCTGCAGCAGAGACCAATNCCACAGCAG 378
QY 597 CTTGCCAATCT 656
Db 377 CTTGCCAATCT 318
QY 657 AACACGCCCAATGTCACAGCAGCGCCCTCACAGTGTGTGCCCGCCCGCTCTACAGGAGC 716
Db 317 AACACGCCCAATGTCACAGCAGCGCCCTCACAGTGTGTGCCCGCCCGCTCTACAGGAGC 258
QY 717 ATGAGGACTGGCAGGCTCCCTCTGGAGAGCCCATCTCTGCTTGGCGTCTACAGGAGC 776
Db 257 ATGAGGACTGGCAGGCTCCCTCTGGAGAGCCCATCTCTGCTTGGCGTCTACAGGAGC 198
QY 777 ACCCGCGTCTGAGTGAAGGAGATTGGGGGTTTCAGGATAGGGAATGGGAGGTC-AG 835
Db 197 ACCCGCGTCTGAGTGAAGGAGATTGGGGGTTTCAGGATAGGGAATGGGAGGTCNAG 138
QY 836 AGGACGCAAGACGACGACCATGTAGATGAACCGTCCAGAGAGCCACGACGACGAGGA 895

Db 137 AGAGCGCAAGACGACGACCATGTAGAACTGTCAGAGAGCCACGACGACGAGGA 78
QY 896 CTGCAGGCATCAGCGTGCACCTGTTCTGATTTTGGAGTTTCATCAAAATCAGTGTGTTT 955
Db 77 CTGCAGGCATCAGCGTGCACCTGTTCTGATTTTGGAGTTTCATCAAAATCAGTGTGTTT 18
QY 956 GCTGCTCTTGGCCACAAA 972
Db 17 GCTGCTCTTGGCCACAAA 1
RESULT 4
H73373
LOCUS H73373 480 bp mRNA linear EST 31-OCT-1995
DEFINITION YU48f10.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
IMAGE:229387 5', mRNA sequence.
H73373
ACCESSION H73373.1 GI:1047623
VERSION H73373.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 480)
AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,
Chisoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins,
M., Hultman, M., Kucaba, T., Lacy, M., Le, N., Le, N., Mardis, E., Moore,
B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T.,
Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E.,
Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
97044478
PUBMED
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLMNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1102 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 306.
Location/Qualifiers
1..480
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3780483"
/db_xref="taxon:9606"
/clone="IMAGE:229387"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen INFLS"
/note="Organ: Liver and Spleen; Vector: p773D (Pharmacia)
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5' AACTGGAAGTAATTAAGATCTTTTATTTTATTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 100 a 149 c 130 g 100 t 1 others
ORIGIN
Query Match 37.6%; Score 373.2; DB 14; Length 480;

```

Best Local Similarity 92.0%; Pred. No. 3.3e-43;
Matches 439; Conservative 0; Mismatches 29; Indels 9; Gaps 4;

QY 12 GGAACCTAGAGGTTCTCACTGCCGAGCAGAGCCCTACACCACCGAGGCAATGGGCTC 71
Db 4 GAAACTAGAGGTTCTCACTGCCGAGCAGAGCCCTACACCACCGAGGCAATGGGCTC 63
QY 72 CTTGGGCTTTCTGTTGGCCGTGCTGGTGGCCAGCAGCTTCTCCAAAGCAGCGGAGGAA 131
Db 64 CTTGGGCTTTCTGTTGGCCGTGCTGGTGGCCAGCAGCTTCTCCAAAGCAGCGGAGGAA 123
QY 132 GAATATACCCCTGGTCTCAATGCTTCAAGTCTGGAAGTTCTTCCAAAGCGCGC 191
Db 124 GAATATACCCCTGGTCTCAATGCTTCAAGTCTGGAAGTTCTTCCAAAGCGCGC 183
QY 192 TGGTGTCTATAACCTGTGTGACCCAGCCAGCCAGCCAGCCAGCTATTCCTCTGT 251
Db 184 TGGTGTCTATAACCTGTGTGACCCAGCCAGCCAGCCAGCCAGCTATTCCTCTGT 243
QY 252 GGAACCAAGACATCAAGTGGCCAAAGAGTGGTGAAGACCCAGCGCGCTCTCTC 311
Db 244 GGAACCAAGACATCAAGTGGCCAAAGAGTGGTGAAGACCCAGCGCGCTCTCTC 303
QY 312 AACCTCAAGTCACACTCAAGTCAGTCCAGACTGCTTCACTTCTGCGGCG 367
Db 304 AACCTCAAGTCACACTCAAGTCAGTCCAGACTGCTTCACTTCTGCGGCG 363
QY 368 GTCTCCACCTCAGTCCCATGT-GCACAGTGCAGGCT--ACAGATGCACTGGAGC 423
Db 364 TTCTCCACCTCAGTCCCATGTGGGACAGTGCAGGTTACAGATGGCAATGGGAGC 423
QY 424 TGTGGTCCAAAGCCAGT-TCAGAGTCCGGCCAACTTCACTCTGACAGGACAGAGG 479
Db 424 TGTGGTCCAAAGCCAGT-TCAGAGTCCGGCCAACTTCACTTTCAGGACAGAGG 480

RESULT 5
H58326
LOCUS
DEFINITION
Yr25c07.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
IMAGE:206316 5', mRNA sequence.
ACCESSION
H58326
VERSION
H58326.1 GI:1011158
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman,
M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
Rifkin, L., Rohlfs, T., Soares, M., Tan, F., Trevaskis, E., Waterston,
R., Williamson, A., Wohldmann, P. and Wilson, R.
The WashU-Merck EST Project
Unpublished
CONTACT: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 912
High quality sequence stops: 251
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 912 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 251.
Location/Qualifiers
1..418
/organism="Homo sapiens"
/mol_type="mRNA"

FEATURES
source
H74302
LOCUS
DEFINITION
Yus6c11.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
IMAGE:230132 5', mRNA sequence.
ACCESSION
H74302
VERSION
H74302.1 GI:1047713
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiappelli, B.,
Chisoe, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W., Hawkins,
M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore,
B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfs, T.,
Schellenberg, K., Soares, M.B., Tan, F., Thierly-Weg, J., Trevaskis, E.,
Underwood, K., Wohldmann, P., Waterston, R., Wilson, R. and Marra, M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)

/db_xref="GDB:3775447"
/db_xref="taxon:9606"
/clone="IMAGE:206316"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen INFLS"
/note="Organ: Liver and Spleen; Vector: p773D (Pharmacia)
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;
1st strand cDNA was primed with a Pac I oligo (dT) primer
[5', AACTGGAGATTAATTAAGATCTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 89 a 130 c 119 g 76 t
ORIGIN
Query Match 35.7%; Score 354; DB 14; Length 418;
Best Local Similarity 96.1%; Pred. No. 1.7e-40;
Matches 393; Conservative 0; Mismatches 13; Indels 3; Gaps 3;

QY 279 AAGTGTGTGAGACCCAGAGCGGCTCTTCAACCTCAACTCACTCAAGTCCAGT 338
Db 3 AGTGGGTGAAGACCCAGAGCGGCTCTTCAACCTCAACTCACTCAAGTCCAGT 62
QY 339 CCAGACTGCTCACTCACTTCTGCGGCGGCTCTTCAACCTCAACTCACTCAAGTCCAGT 398
Db 63 CCAGACTGCTCACTCACTTCTGCGGCGGCTCTTCAACCTCAACTCACTCAAGTCCAGT 122
QY 399 GCAGGCTACAGATGCACTGGAGCTGTGTCCTCAAGCCAGTGTCTGAGTCCGGGCAAC 458
Db 123 GCAGGCTACAGATGCACTGGAGCTGTGTCCTCAAGCCAGTGTCTGAGTCCGGGCAAC 192
QY 459 TTCACTCTCAGACAGAGGCGAGCCCGAGGTGGAGATGATTCGCCAGGCGTCTCG 518
Db 183 TTCACTCTCAGAGACAGAGGCGAGG-CCAGGTTGGAGATGATTCGCCAGGCGTCTCG 241
QY 519 GGAGGCCACCTATCACCACAGGCTGTGTCCTCAAGCCAGTGTCTGAGTCCGGGCAAC 578
Db 242 GGAGGCCACCTATCACCACAGGCTGTGTCCTCAAGCCAGTGTCTGAGTCCGGGCAAC 301
QY 579 ABACATGCCACAGGCGCTTCCAACTTCTCTCTCC-TGCCAGCCAGACATCGGACTG 637
Db 302 AGACATGCCACAGGCGCTTCCAACTTCTCTCTCC-TGCCAGCCAGACATCGGACTG 361
QY 638 G-TTCTGTTGCCAGGCTGCNACACAGCCCAATGTCCAGCAGAGCGCCCT 685
Db 362 GTTTTGTGCCAGGCTGCNACACAGCCCAATGTCCAGCAGAGGCGCCCT 410

RESULT 6
H74302
LOCUS
DEFINITION
Yus6c11.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
IMAGE:230132 5', mRNA sequence.
ACCESSION
H74302
VERSION
H74302.1 GI:1047713
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiappelli, B.,
Chisoe, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W., Hawkins,
M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore,
B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfs, T.,
Schellenberg, K., Soares, M.B., Tan, F., Thierly-Weg, J., Trevaskis, E.,
Underwood, K., Wohldmann, P., Waterston, R., Wilson, R. and Marra, M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)

```



```

Query Match      27.9%; Score 277.2; DB 14; Length 379;
Best Local Similarity 98.9%; Pred. No. 1.1e-29;
Matches 279; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 693 GTGCCCCAGGTTGACACAGAGATGGAGGACTGGCAGGTCCTCCCTGGAGAGCCCCATC 752
Db 282 GTGGCCCCAGGTTGACACAGAGATGGAGGACTGGCAGGTCCTCCCTGGAGAGCCCCATC 223

QY 753 TTGCTTTGCCCTCTACAGAGACACCCGCCCTCTGAGTGAAGAGAGCTTTGGGGGGTTC 812
Db 222 TTGCTTTGCCCTCTACAGAGACACCCGCCCTCTGAGTGAAGAGAGCTTTGGGGGGTTC 163

QY 813 AGGATAGGATGGGGAGTCAAGGACGCAAGCAGCAGGCGCATGTAGAATGAACCGTCC 872
Db 162 AGGATAGGATGGGGAGTCAAGGACGCAAGCAGCAGGCGCATGTAGAATGAACGTGCC 103

QY 873 AGAGAGCAAGCACCGCAGAGGACTGCAGGCGCATCAGCGTGCACCTGTTCTGATTTGGAGT 932
Db 102 AGAGAGCAAGCACCGCAGAGGACTGCAGGCGCATCAGCGTGCACCTGTTCTGATTTGGAGT 43

QY 933 TCATGCAAAATGAGTGTGTTTGTAGTCTGCTCTTGCACAAAAA 974
Db 42 TCATGCAAAATGAGTGTGTTTGTAGTCTGCTCTTGCACAAAAA 1

RESULT 12
LOCUS AI438986 450 bp mRNA linear EST 18-MAR-1999
DEFINITION tc84a08.x1 NCI CGAP CLL1 Homo sapiens cDNA clone IMAGE:2072822 3'
similar to contains 'Alu repetitive element; contains element PTR5
repetitive element ;, mRNA sequence.
ACCESSION AI438986
VERSION AI438986.1 GI:4301160
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 450)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapb-remail.nih.gov
Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever,
M.D., Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/image.html at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1792 Std Error: 0.00
Seq primer: -40UP from Gibco.
FEATURES
Location/Qualifiers
1..450
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2072822"
/tissue_type="B-cell, chronic lymphocytic leukemia"
/lab_host="DH10B"
/clone_lib="NCI CGAP CLL1"
polylinker: Vector: pTZ19D-pac (Pharmacia) with a modified
Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'-
TGTACATCTCACTGAGTGGCGGCGGCATTCGTTTCTTTTCTTTTCTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pTZ19 vector.
Library is normalized, and was constructed by Bento

```

```

BASE COUNT      Soares and M.Patima Bonaldo."
ORIGIN          85 a 144 c 109 g 112 t

Query Match      27.6%; Score 274.2; DB 9; Length 450;
Best Local Similarity 97.2%; Pred. No. 2.6e-29;
Matches 279; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 688 CAGTGGTGGCCCCAGGTGTGACACAGAGATGGAGGACTGGCAGGTCCTCCCTGGAGAGCC 747
Db 287 CACTGCTTTGTCAAGGTGGTGCACACAGAGATGGAGGACTGGCAGGTCCTCCCTGGAGAGCC 228

QY 748 CCATCCTTGGCTTGGCCCTCTACAGGAGCACCAGGCTCTGAGTGAAGAGAGGTTTGGGG 807
Db 227 CCATCCTTGGCTTGGCCCTCTACAGGAGCGCCCGCTCTGAGTGAAGAGAGGTTTGGGG 168

QY 808 GGTTCAGATAGGAATGGGAGGTCAAGAGCGCAAGCAGCAGGCGCATGTAGAATGAAC 867
Db 167 GGTTCAGATAGGAATGGGAGGTCAAGAGCGCAAGCAGCAGGCGCATGTAGAATGAAC 108

QY 868 CGTCCAGAGAGCCAAAGCAGCGGAGGACGAGGCGCATCAGCGTGCACCTGTTCTGTTT 927
Db 107 CGTCCAGAGAGCCAAAGCAGCGGAGGACGAGGCGCATCAGCGTGCACCTGTTCTGTTT 48

QY 928 GGAGTTTCATGCAAAATGAGTGTGTTTGTAGTCTGCTCTTGCACAAAAA 974
Db 47 GGAGTTTCATGCAAAATGAGTGTGTTTGTAGTCTGCTCTTGCACAAAAA 1

RESULT 13
LOCUS BY718164 553 bp mRNA linear EST 17-DEC-2002
DEFINITION BY718164 RIKEN full-length enriched, 13 days embryo male testis Mus
musculus cDNA clone 6030468B19 5', mRNA sequence.
ACCESSION BY718164.1 GI:27131281
VERSION BY718164
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 553)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H.,
Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,
Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A.,
Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S.,
Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V., Chothia, C., Corbani,
J. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest,
A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J.,
Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M.,
King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons,
P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Mikl,
H., Negashima, T., Numata, K., Okido, T., Pavan, W. J., Pertea, G.,
Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D.,
Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring,
B. Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C. A., Setou,
M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale,
R. D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y.,
Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa,
M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura,
M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Imotani, K., Ishii,
Y., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Kawai, J., Aizawa, K.,
Y. Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata,
K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander,
E. S., Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
TITLE Nature 420, 563-573 (2002)
JOURNAL Nature 420, 563-573 (2002)
MEDLINE 22354683
PUBMED 12466851

```


Db 400 TG 401

RESULT 23
T87491

LOCUS
DEFINITION
T87491.1 GI:715843
324 bp mRNA linear EST 17-MAR-1995
Y85903.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
IMAGE:115060 5', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)

REFERENCE
AUTHORS
Hallier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaakis,E., Waterston
R., Williamson,A., Wohldmann,P. and Wilson,R.
The WashU-Merck EST Project
Unpublished

TITLE
JOURNAL
COMMENT
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1048
High quality sequence stops: 179 Source: IMAGE Consortium, LMLL
This clone is available royalty-free through LMLL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1048 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 179.

FEATURES
source
1. .324
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:470677"
/db_xref="taxon:9606"
/clone="IMAGE:115060"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen INFLS"
/note="Organ: Liver and Spleen; Vector: p773D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5' AACTGGAAGAATTAATTAAGATCTTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT
ORIGIN
70 a 99 c 77 g 78 t

Query Match 19.4%; Score 192.2; DB 14; Length 324;
Best Local Similarity 97.7%; Pred. No. 9.4e-18;
Matches 216; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 131 AGAAATTACCCCTGTGCTCTCCATTGCGTACAAAGTCCTGGAAGTTTCCCAAGGCCG 190
|||||

Db 75 AGAAATTACCCCTGTGCTCTCCATTGCGTACAAAGTCCTGGAAGTTTCCCAAGGCCG 134
|||||

QY 191 CTGGGTGCTCATAACTGCTGTGCACCCAGCCACACCGCCCATCACTATTTCCTCTG 250
|||||

Db 135 CTGGGTGCTCATAACTGCTGTGCACCCAGCCACACCGCCCATCACTATTTCCTCTG 194
|||||

QY 251 TGGAAACAAGAAATCAAGT-GGCAAGAAGTGGTGAAGACCCACGAGCGGCTCC- 308
|||||

Db 195 TGGAAACAAGAAATCAAGTGGCGCAAGAAGTTGTGAAGACCCACGAGCGGCTCT 254
|||||

126	QY	GAGGAGAAATTACCCCTGTGTGCTCTCCATTCGCTACAAAGTCTCGAAGTTTCCCCAA	185
74	Db	CAGAAGGCA-----TCACCATTCGCTACAAAGTACTGAAGTTTATCCCAA	120
186	QY	GGCGCTGGTGCTCATAACTGCTGTGTCACCCCCAGGCCACCGGCCCATCACTATTCC	245
121	Db	AGCCGGA--GGTGCTTATAACCTGCGATGCCCTCAGCGCTCCAGGCCCATCACATACT	179
246	QY	CTCTGTGGAA--CCAGAACAATCAAGTGTGGCCAGAAAGTGTGGAAGACCCACGAGCCGG	303
180	Db	CTCTCGGCTAGCCGATGTCTCTCGTGGCGAAAAAGTTGTGCATGCTCGTGGCCGG	239
304	QY	CTCTCTTCAACTCAACGTCACTCAAGTCCAGTCCAGACCTGCTCACTACTTTGCGC	363
240	Db	CCTCTCTCAACATCAATATCAACATCAAGTCCAGCCAGACCTGCTCACCTACTCTCTGCG	299
364	QY	GGCGCTCTTCCACCTCAGCTGCCATGTGGACAGTGGCCAGGCTACAGATGCATGGGAGC	423
300	Db	AGGCAACCTTCGAACCTTGGACCTTAGGACCCAGCAGGAGCTCCAGATGTACAGGAAC	359
424	QY	TGTGTGTCCAAAGCCAGTGTCTGAGCTGCGGGGCCAACTTCACTCTGCAGGAACAGAGGGGCGAG	483
360	Db	TGTGGGCTAAG--CAGTGTCTCAGCTCAGAGGCTGACTCTGCTCTACGGCATGGGAGATCGG	418
484	QY	GCCCAAGGTTGAGATGATCTGCGCAGGCGTCTCTGGGGAGGCCCACTTAT	532
419	Db	GGCCCACTGTGAGCTGTCTCTGCTGGCATCTCTCAGGCAGCCCCCCCAT	467

RESULT	25
N86939	
LOCUS	N86939
DEFINITION	Cl498F Human fetal heart, Lambda ZAP Express Homo sapiens cDNA
ACCESSION	N86939
VERSION	N86939.1
KEYWORDS	GI:1440141
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
	1. (bases 1 to 218)
	Liew,C.C.
	cDNAs from fetal heart (1996)

BASE COUNT 50 a 48 c 71 g 49 t
ORIGIN
Query Match 17.7%; Score 175.6; DB 14; Length 218;
Best Local Similarity 97.1%; Pred. No. 2.4e-15;
Matches 200; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

[illegible]

LOCUS	BY212485	364 bp	mRNA	linear	EST 10-DEC-2002
DEFINITION	BY212485 RIKEN full-length enriched, activated spleen Mus musculus cDNA clone F830016F14 5', mRNA sequence.				
ACCESSION	BY212485				
VERSION	BY212485.1	GI:26393130			
KEYWORDS	EST.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Eumalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				

TITLE	JOURNAL	MEDLINE	PUBMED	COMMENT
-------	---------	---------	--------	---------

LOCUS BI347003 376395 MARC 2P1G Sus scrofa cDNA 5', mRNA linear EST 30-JUL-2001
 DEFINITION 376395 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
 ACCESSION BI347003
 VERSION BI347003.1 GI:15040292
 KEYWORDS EST.
 SOURCE Sus scrofa (pig)
 ORGANISM Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 REFERENCE 1 (bases 1 to 582)
 AUTHORS Fahrénkrug, S.C., Smith, T.P.L., Freking, B.A., Cho, J., White, J., Vallet, J., Wise, T., Rohrer, G.A., Perce, G., Sultana, R., Quackenbush, J., and Keele, J.W.
 TITLE Porcine gene discovery by normalized cDNA-library sequencing and EST cluster assembly
 JOURNAL Mamm. Genome 13 (8), 475-478 (2002)
 MEDLINE 22213789
 PUBMED 12226715
 COMMENT Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.
 PCR Primers
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCCTCAGCAGCG
 Plate: 130 row: N column: 8
 Seq primer: ATTATGGTGACACTATAG.
 Location/Qualifiers
 1..582
 /organism="Sus scrofa"
 /mol_type="mRNA"
 /db_xref="taxon:9823"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_lib="MARC 2P1G"
 /note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI; Library made from pooled tissue from testis, ovary, endometrium, hypothalamus, pituitary, and placenta."
 BASE COUNT 122 a 191 c 166 g 103 t

FEATURES

source

Query Match 15.5%; Score 154; DB 12; Length 582;
 Best Local Similarity 75.2%; Pred. No. 1.7e-12;
 Matches 206; Conservative 0; Mismatches 65; Indels 3; Gaps 1;
 QY 433 AGCCAGTGTCTGAGCTGGGGCCAACTTCACTCTGCAGGACAGAGG---GGCAGGCCCA 489
 Db 132 AGCCCGTGTCCAGCTGCTGTTTAACTTTCATCTCTGCGAGGAGCCCGTGGGCGCCCT 191
 QY 490 GGGTGGAGATGATCTGCGAGCGTCTCTGGGAGCCACCTATCACCAACAGCTGATCG 549
 Db 192 GGGTGAAGGTGTGCTGCTGGGATCTCTGGGAGCCACCTACCTACAGTCTTGTGCG 251
 QY 550 GGAAGGATGGGAGGTCTCCACCTCGAGCAGAGACCATGCCAGGCGCTGCCAACTTCT 609
 Db 252 GGAGGGACAGTGGCAGCTTACATGAGCAGACACAGAACTACAGGAGCTGCCAACTTCT 311
 QY 610 CTTCTCTCGAGCCAGACATCGGACTGGTCTGTGTGTCAGGCTGCAACACGCCAATG 669
 Db 312 CTTCTCTCGAGCCAGACATCGGACTGGTCTGTGTGTCAGGCTGCAACACGCCAATG 669
 QY 670 TCCAGCAGCGCCCTCACTGGTGGCCCGCAGG 703
 Db 372 CCCAGCAGTGGCCCTCACTGGTGGCCCGCAGG 405

RESULT 29
 CH051335

LOCUS CB051335 443 bp mRNA linear EST 17-JAN-2003
 DEFINITION NISC_gj23h11.y1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:3289604
 ACCESSION CB051335
 VERSION CB051335.1 GI:27789622
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 443)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 CDNA Library Preparation:
 DNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC)
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
 Plate: L1AM8055 row: P column: 21
 Seq primer: M13RP1 reverse primer (ABI).
 Location/Qualifiers
 1..443
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3289604"
 /sex="male"
 /dev stage="adult"
 /lab_host="DH10B"
 /clone_lib="NCI_CGAP_Pr28"
 /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI_CGAP_Pr22 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 985608-986759, 1101192-1101959, and 1217928-1220615). Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 129 a 112 c 111 g 91 t

ORIGIN
 Query Match 15.5%; Score 153.8; DB 14; Length 443;
 Best Local Similarity 98.7%; Pred. No. 2e-12;
 Matches 155; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 833 CAGAGGACGCAAGACGACGCGCATGTAGATGAACCTCCAGAGAGCAAGCAGCGAGA 892
 Db 250 CAGAGGACGCAAGACGACGCGCATGTAGATGAACCTCCAGAGAGCAAGCAGCGAGA 309
 QY 893 GGACTCGAGCCCATCAGCGTGCACCTGTCGTATTGGAGTTCATGCAAAATGAGTGTGT 952
 Db 310 GGACTCGAGCCCATCAGCATGCATCTTCGTATTGGAGTTCATGCAAAATGAGTGTGT 369
 QY 953 TTAGCTGCTTTCGCCACAAAAAATAAAAAA 989
 Db 370 TTAGCTGCTTTCGCCACAAAAAATAAAAAA 406

RESULT 30
 CB051334/c

LOCUS

CB051334

DEFINITION

NISC_gj23h11.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:3289604

ACCESSION

CH051334

VERSION

CB051334.1 GI:27789621

KEYWORDS

EST.

SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 540)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
cDNA Library Preparation:
cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC)
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: infoimage.llnl.gov
Plate: LLAM8055 row: P column: 21
Seq primer: -21M13 forward primer (ABI).
FEATURES Location/Qualifiers
1..540
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3289604"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/notes="Organ: Prostate; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI CGAP Pr22 was prepared, and as this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonoids 985608-986759, 1101192-1101959, and 1217928-1220615). Subtraction by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 126 a 113 c 128 g 173 t
ORIGIN
Query Match 15.5%; Score 153.8; DB 14; Length 540;
Best Local Similarity 98.7%; Pred. No. 1.8e-12; Indels 0; Gaps 0;
Matches 155; Conservative 0; Mismatches 2;
QY 833 CAGAGCAGCAAGCAGCAGCCATGTAGAAATGAACCGTCCAGAGAGCCCAAGCAGCAGA 892
Db 305 CAGAGCAGCAAGCAGCAGCCATGTAGAAATGAACCGTCCAGAGAGCCCAAGCAGCAGA 246
QY 893 GGACTCAGGCCATCAGCGTGCACCTGTTGTTATTTGGAGTTCTATGCAAAATGAGTGTGT 952
Db 245 GGACTCAGGCCATCAGCGTGCACCTGTTGTTATTTGGAGTTCTATGCAAAATGAGTGTGT 186
QY 953 TTAGCTGCTCTTCCCAAAAAA 969
Db 185 TTAGCTGCTCTTCCCAAAAAA 149
RESULT 31
LOCUS N88474
DEFINITION X3475F Human fetal heart, Lambda ZAP Express Homo sapiens cDNA clone K3475 5', mRNA sequence.
ACCESSION N88474
VERSION N88474.1 GI:1441676
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 194)
AUTHORS Liew, C.C.

SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 382)
AUTHORS Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Imotani, K., Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akanira, S., Tanaka, T., Tomaru, A., Toya, T., Watabiki, A., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al. 2001)
JOURNAL Unpublished
COMMENT Contact: Yoshihide Hayashizaki

SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 382)
AUTHORS Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Imotani, K., Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akanira, S., Tanaka, T., Tomaru, A., Toya, T., Watabiki, A., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al. 2001)
JOURNAL Unpublished
COMMENT Contact: Yoshihide Hayashizaki

SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 382)
AUTHORS Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Imotani, K., Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akanira, S., Tanaka, T., Tomaru, A., Toya, T., Watabiki, A., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al. 2001)
JOURNAL Unpublished
COMMENT Contact: Yoshihide Hayashizaki

TITLE cDNAs from fetal heart (1996)
JOURNAL Unpublished
COMMENT Contact: Liew CC
Brigham and Women's Hospital
Harvard Medical School
75 Francis St. Boston, MA 02115, USA
Tel: 6177328915
Fax: 6179750995
Email: cliew@rics.bwh.harvard.edu
Seq primer: GAAATTACCCCTCACTAAAGG.
FEATURES Location/Qualifiers
1..194
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="K3475"
/lab_host="E. coli XL1-Blue"
/notes="Vector: Lambda ZAP Express; Site 1: EcoRI; Site 2: XhoI; mRNA was purified from human fetal hearts (8-10 weeks). cDNA was synthesized using a XhoI-Oligo dT adaptor-primer. EcoRI adaptors were ligated, followed by digestion with XhoI, for directional cloning into predigested lambda ZAP Express."
BASE COUNT 51 a 43 c 63 g 37 t
ORIGIN

Query Match 14.6%; Score 145.2; DB 14; Length 194;
Best Local Similarity 96.5%; Pred. No. 4.6e-11;
Matches 191; Conservative 0; Mismatches 3; Indels 4; Gaps 4;
QY 750 ATCCCTGCTCCGCTCTACAGGAGCACCCTCTGAGTGAAGAGAGTTGGGGG 809
Db 1 ATCCCTGCTCCGCTCTACAGGAGCACCCTCTGAGTGAAGAGAGTTGGGGG 59
QY 810 TTCAGATAGGAATGGGGAGTTCAGAGGACCAAGCAGCAGCCATGTAGATGAACG 869
Db 60 TTCAGATAGGAATGGGGAGTTCAGAGGACCAAGCAGCAGCCATGTAGATGAACG 119
QY 870 TCCAGAGAGCCCAAGCAGCAGGAGTTCAGGAGCCATCAGCGTCTGCTGTTATTTGG 929
Db 120 T-CAGAGAGCCCAAGCAGCAGGAGTTCAGGAGCCATCAGCGTCTGCTGTTATTTGG 176
QY 930 ACTTATGCAAAATGAGT 947
Db 177 ACTTATGCAAAATGAGT 194

RESULT 32
LOCUS BB842514
DEFINITION BB842514 RIKEN full-length enriched, 6 days neonate spleen Mus musculus cDNA clone F430005P21 5', mRNA sequence.
ACCESSION BB842514
VERSION BB842514.1 GI:17043245
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 382)
AUTHORS Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Imotani, K., Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akanira, S., Tanaka, T., Tomaru, A., Toya, T., Watabiki, A., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al. 2001)
JOURNAL Unpublished
COMMENT Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@gsr.riken.go.jp/
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.
e mouse tissues.

FEATURES Location/Qualifiers
Source 1..382

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="F43005P21"
/tissue_type="spleen"
/dev_stage="6 days neonate"
/clone_lib="RIKEN full-length enriched, 6 days neonate spleen"

BASE COUNT 87 a 122 c 99 g 74 t

ORIGIN
Query Match 14.4%; Score 143; DB 10; Length 382;
Best Local Similarity 69.8%; Pred. No. 7e-11;
Matches 213; Conservative 0; Mismatches 80; Indels 12; Gaps 1;

QY 44 GCCCTACACCCAGGAGGATGGGGCTCCCTGGGCTGTCTGTGGCGGTGGTGGCTGC 103
DB 90 GCCCAACCTCAGAGGATGGGGCTCCCTTCAGTCTCTCTTTGCCATGGTGGTGC 149
QY 104 CAGCAGCTTCTCAAGGACGGGAGGAAGAAATPACCCCTGTGGTCTCCATTGCTACAA 163
DB 150 CTGTGGCTTCTCAGAGGAGCAGACAGAGGCATCA-----CCATTGCTACAA 197
QY 164 AGTCTCGAGTCTTCCCAAGGCGCTGGTGTCTATACTGTCTGTGCAACCCAGCC 223
DB 198 AGTACTGGAAGTTATCCCAAGCCGGAGGGTGTCTTATACTGTGCAACCCAGCC 257
QY 224 ACCACCCGCCATCACCTATTCCCTCTGTGAACCAAGAAATCAATCAAGTGGCCAGAGGT 283
DB 258 GTCCAGCCCATCACATCTCTCTCTGGTGTAGCGAGGTATCTGTGGCAAAAAGGT 317
QY 284 GGTGAAGACCCAGGCGGCTCTTCAACCTCAACGTCACACTCAAGTCCAGTCCAGA 343
DB 318 TGTGCATGACTCCGTGGCGGCTCTCTTCAACATCATCAATCAATCAAGTCCAGCCAGA 377
QY 344 CCTGC 348
DB 378 CCTGC 382

RESULT 33
BB843270

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

TITLE

JOURNAL
COMMENT

Unpublished

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsr.riken.go.jp/
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.
e mouse tissues.

FEATURES Location/Qualifiers

Source 1..370

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="F430010H21"
/tissue_type="spleen"
/dev_stage="6 days neonate"
/clone_lib="RIKEN full-length enriched, 6 days neonate spleen"

BASE COUNT 85 a 116 c 96 g 73 t

ORIGIN

Query Match 13.4%; Score 132.6; DB 10; Length 370;
Best Local Similarity 68.9%; Pred. No. 2e-09;
Matches 202; Conservative 0; Mismatches 79; Indels 12; Gaps 1;

QY 44 GCCCTACACCCAGGAGGATGGGGCTCCCTGGGCTGTCTGTGGCGGTGGTGC 103

DB 90 GCCCAACCTCAGAGGATGGGGCTCCCTTCAGTCTCTCTTTGCCATGGTGGTGC 149

```

QY 104 CAGCAGCTTCTCAAGCAGCGGAGGAGAAATTACCCCTGTGGTCTCCATTCCCTACAA 163
Db 150 CTGTGGCTTCTCAGAGAGCAGACAGAGGCATCA-----CCATTGCCCTACAA 197
QY 164 AGTCTCTGGAAGTTTTCCTCCAAAGCGCGCTGGGTCTCATAACTGTGTGTCACCCAGCC 223
Db 198 AGTACTGGAAGTTTATCCCAAGCGGAGGGTGTCTTAACCTGCGATGCCCTCAGGC 257
QY 224 ACCACGCCCATCACCTATTCCTCTGTGTGGAACAAGAACATCAAGTGGCCCAAGAGGT 283
Db 258 GTCCAGGCCCATCACATACCTCTCTCTGGCTAGCCGAGGTATCTCTGTGGCAAAAAGGT 317
QY 284 GTTGAAGACCCAGCAGCGCGCTCTCTCACTCAAGTCACACTCAAGTCCA 336
Db 318 TGTGATGACTCGTCCGCGGCTCTCTCAACATCAATATCACCATCAAGTCCA 370

RESULT 34
BF443482
LOCUS 260980 MARC 2P1G Sus scrofa cDNA 5', mRNA linear EST 01-DEC-2000
DEFINITION
ACCESSION BF443482
VERSION BF443482.1 GI:11503574
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE
AUTHORS Fahrrenkrug,S.C., Smith,T.P.L., Freking,B.A., Cho,J., White,J.,
Vallet,J., Wise,T., Rohrer,G.A., Perlea,G., Sultana,R., Quackenbush
J. and Keele,J.W.
TITLE Porcine gene discovery by normalized cDNA-library sequencing and
EST cluster assembly
JOURNAL Mamm. Genome 13 (8), 475-478 (2002)
MEDLINE 2213789
PUBMED 12226715
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTCCAGTCACGACG
Plate: 93 row: K column: 24
Seq primer: ATTTAGTGACACTATAG.
FEATURES
source
1..420
Location/Qualifiers
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 2P1G"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT 99 a 109 c 130 g 82 t
ORIGIN
Query Match 9.4%; Score 93.2; DB 10; Length 420;
Best Local Similarity 71.1%; Pred. No. 0.00064;
Matches 138; Conservative 0; Mismatches 53; Indels 3; Gaps 1;

QY 428 GTCCAGCCAGTGTCTGAGCTCGGGCCCACTTCACTTCGACGACAGAGG---GGCAGG 484
Db 223 GCCAGAGCCCGTGTCCAGCTGTGTGTTAACTTCATCTCTGCTGGACGAGGCCGCTCGGG 282

```

```

QY 485 CCCCAGGGTGGAGATGATTCGCAGCGTCTCTGGCGAGCCCACTATCAACACAGCCT 544
Db 283 CCCTTGGGTGAAGGTTCGTGTGTGGGCATCTCTGGGCGAGCCCACTACCTACAGTCT 342
QY 545 GATCGGGAAGATGGCAGGTCCACTGCGACGAGACCATGCGACAGGAGCTGCCAA 604
Db 343 TGTGCGGAGGAGCAGTGGCCACCTATATGAAGCAGACACAGAACTACAGGAAGCTGCCAA 402
QY 605 CTTCTCCTTCCTGC 618
Db 403 CTTCTCCTTCCTGC 416

RESULT 35
BF442092
LOCUS 258654 MARC 2P1G Sus scrofa cDNA 5', mRNA linear EST 01-DEC-2000
DEFINITION
ACCESSION BF442092
VERSION BF442092.1 GI:11502184
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE
AUTHORS Fahrrenkrug,S.C., Smith,T.P.L., Freking,B.A., Cho,J., White,J.,
Vallet,J., Wise,T., Rohrer,G.A., Perlea,G., Sultana,R., Quackenbush
J. and Keele,J.W.
TITLE Porcine gene discovery by normalized cDNA-library sequencing and
EST cluster assembly
JOURNAL Mamm. Genome 13 (8), 475-478 (2002)
MEDLINE 2213789
PUBMED 12226715
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTCCAGTCACGACG
Plate: 88 row: H column: 8
Seq primer: ATTTAGTGACACTATAG.
FEATURES
source
1..503
Location/Qualifiers
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 2P1G"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT 104 a 160 c 160 g 79 t
ORIGIN
Query Match 9.3%; Score 92.6; DB 10; Length 503;
Best Local Similarity 70.8%; Pred. No. 0.00072;
Matches 138; Conservative 0; Mismatches 54; Indels 3; Gaps 1;

QY 433 AGCCAGTCTCTGAGCTGGCGCAACTTCACTTCGACGACAGAGG---GGCAGGCCCA 489
Db 309 AGCCCGTCCCGAGCTGCTGTGTTAACTTCATCTCTGCTGGACGAGGCCGCTT 368
QY 490 GGGTGGAGATGATTCGCCAGCGCTCTCGGGCAGCCCACTATCAACACAGCCTGATCG 549
Db 369 GGGTGAAGGTCTCGTGTGGGCATCTCTGGGCGAGCCCACTACCTACAGTCTTGTGCG 428

```



```

Db      284 AGGTCAGAGCATAAAAGTAGCAGGTATGAGTGAACAGGCTAGAACTACTGTATA 225
QY      887 GGCAGAGGACTGAGGCGATCAGCGTGCACCTGTCGATTGTTGGAGTTTCATGCAAAATGAG 946
Db      224 ACATGAGGAGCTGAGTGAATAAATTTGACT---GTATGTGGAGTTTCATGCTAAATGAG 169
QY      947 TGTGTTTTAGCTGCTCTTGGCACAAAAAATAAAAAAAAAAAAAAAAAAAAA 991
Db      168 TAGATTTTAGCTGCTCTTGGCACAGAAACAAAAATCTAAGGGA 124

RESULT 38
AW794912/c
LOCUS      AW794912                466 bp    mRNA    linear    EST 16-MAY-2000
DEFINITION RC6-UM0015-210200-011-D07 UM0015 Homo sapiens cDNA, mRNA sequence.
ACCESSION  AW794912
VERSION     AW794912.1 GI:7846691
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 466)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,F., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=kt2=RC6-UM0015-210
200-011-D07&t3=2000-02-21&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 14
High quality sequence stop: 462.
FEATURES
Location/Qualifiers
1..466
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="UM0015"
/notes="Organ: uterus; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORSTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT      113 a      98 c      84 g      171 t
ORIGIN
Query Match      5.3%; Score 52.8; DB 9; Length 466;
Best Local Similarity 58.5%; Pred. No. 2.9e+02;
Matches 110; Conservative 0; Mismatches 77; Indels 1; Gaps 1;

QY      804 GGGGGTTCAGGATAGGAATGGGAGGTGAGAGGCGAAGCAGCGCCATGTAGAT 863

```

```

Db      461 GGGTGGAAAGAAAATGGGGAGACGAAAGTCAGAGGATGTAAGTAGCAGACACATAGAA 402
QY      864 GAACCGTCCAGAGAGCAGCAGCAGGAGGAGTGGAGCCATCAGCGTGCACACTGTTGCT 923
Db      401 GATAAAGTCTAGCTTTCTTATGTGCAACATGAGGAGCTACAGATAAATAATATAGTGCT-GT 343
QY      924 ATTTGGAGTTTCATGCAAAATGAGTGTGTTTGTAGCTGCTTGGCCACAAAAAATAAAAA 983
Db      342 ATCTGGATTCATGATAAATGAGTAGATTTTGTGCGTCAGCCACAAAAAGCAAAATAA 283
QY      984 AAAAAAA 991
Db      282 AAACAAA 275

RESULT 39
HSM077193/c
ID      HSM077193 standard; RNA; EST; 493 BP.
XX      AC BX487006;
XX      SV BX487006.1
XX      DT 09-MAY-2003 (Rel. 75, Created)
XX      DT 09-MAY-2003 (Rel. 75, Last updated, Version 1)
DE      Homo sapiens mRNA; EST DKFZp686L01254_r1 (from clone DKFZp686L01254)
XX      KW EST; expressed sequence tag.
XX      OS Homo sapiens (human)
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC      Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX      RN [1]
RP      1-493
RA      Koehrer K., Beyer A., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,
RA      Han M., Wiemann S.;
RT      Submitted (07-MAY-2003) to the EMBL/GenBank/DBJ databases.
RL      MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY;
XX      CC This is the 5' sequence of the clone insert
CC      Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
CC      Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
CC      sequenced by BMFZ (Biomedical Research Center at the
CC      Heinrich-Heine-University, Dueseldorf/Germany) within the cDNA
CC      sequencing consortium of the German Genome Project.
CC      No si sequence available.
CC      This clone (DKFZp686L01254) is available at the RZPD in Berlin.
CC      Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6,
CC      14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de
XX      FH Key Location/Qualifiers
FH      source
FH      1..493
FT      /db_xref="taxon:9606"
FT      /mol_type="mRNA"
FT      /organism="Homo sapiens"
FT      /clone="DKFZp686L01254"
FT      /clone_lib="586 (synonym: hlcc3). Vector pSport1_Sfi; host
FT      DH10B; sites SfiIA + SfiIB"
FT      /dev_stage="adult"
FT      /tissue_type="cDNA-collection"
XX      SQ Sequence 493 BP; 125 A; 110 C; 62 G; 196 T; 0 other;

Query Match      5.3%; Score 52.6; DB 2; Length 493;
Best Local Similarity 62.0%; Pred. No. 3e+02;
Matches 101; Conservative 0; Mismatches 59; Indels 3; Gaps 1;

QY      820 GGAATGGGAGGTGAGGAGCGAAGCAGCGCCATGTAGATGAACCGTCAGAGAGC 879

```

```

Db      174 GGGAGATAGTGGAGGATACAAAGTAGCAGATATGTAGATGAAACAAGTC---TAGA 118
QY      880 CAAGCAGCGCAGAGACTGCAGGCCATCAGCGTGTCACCTGTTTCGTAATTTGGAGTTCATGCA 939
Db      117 GAAATTAATGTACATGAGGACCTCTAGGTAACACAATTGTAITGTATGTGGTAATCATGCT 58
QY      940 AAATGAGTGTGTTTACGTCTCTTCGCCACAAAAAAGAAAAA 982
Db      57 AAAGGAGTACAAATTAGTCTCTTGCACAAAAAACAAGAAA 15

RESULT 40
BI063151/c
LOCUS   BI063151          606 bp      mRNA      linear      EST 15-JUN-2001
DEFINITION IL3-UT0117-300301-512-B06 UT0117 Homo sapiens cDNA, mRNA sequence.
ACCESSION BI063151
VERSION   BI063151.1  GI:14470678
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 606)
Dias Neto,E., Garcia Correa,R., Veriovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE    Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL  Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE  20202663
PUBMED   10737800
COMMENT  Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l1=IL3&t2=IL3-UT0117-
300301-512-B06&t3=2001-03-30&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 502.
FEATURES             Location/Qualifiers
     source            1..606
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="UT0117"
                     /note="Organ: uterus_tumor; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT  124 a 167 c 143 g 172 t
ORIGIN
Query Match          5.2%; Score 51.2; DB 12; Length 606;
Best Local Similarity 62.8%; Pred. No. 4.3e+02;
Matches 113; Conservative 0; Mismatches 63; Indels 4; Gaps 2;

QY      813 AGGATAGGAAATGGGAGGTGACAGGACGAAAGCAGCGCCATGTAGAAATGACCGTCC 872
Db      540 AGGAAACGGAGATGTAGTGAAGACGCGAGAGGCGAGATATGCAAGATGCATCGCC 481

```

```

QY      873 AGAGAGCAAGCAGCGCAGAGGACTGCAGGCCATCAGCGTCACCTGTCGTAATTTGGAGT 932
Db      480 AGAG-GCCTAATGTGCATGAGGACTCTCTAGTCAACAACATTGA---ATTATACCGGGAT 425
QY      933 TCATGCAAAATGAGTGTGTTTACGTCTCTTGCACAAAAAAGAAAAAAGAAAAA 992
Db      424 TTCTGCTAAGACAGCTAGATTTTAGGAGCTCTTGCCATATTGAAAAAAGAAAAAGGTAA 365

```

```

RESULT 41
BY560588
LOCUS   BY560588          443 bp      mRNA      linear      EST 15-DEC-2002
DEFINITION BY560588 RIKEN full-length enriched, activated spleen Mus musculus
cDNA clone F830028X04 3', mRNA sequence.
ACCESSION BY560588
VERSION   BY560588.1  GI:26895770
KEYWORDS EST.
SOURCE   Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 443)
Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamakata,I., Kiyosawa,H.,
Yagi,K., Tomaru,X., Hasegawa,Y., Nogami,A., Yoshida,C., Hume,D.A.,
Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A.,
Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batalov,S.,
Beisel,K.W., Blake,J.A., Bradt,D., Brusci,V., Chothia,C., Corbani
,L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest
,A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A.,
Gough,J., Grammond,S., Gustinchin,S., Hirokawa,N., Jackson,I.J.,
Jarvis,E.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kedzierski,R.M.,
King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons
,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki
,H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Pertea,G.,
Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D.,
Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring
,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou
,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale
,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y.,
Watanabe,Y., Wells,C., Wilming,L.G., Wyshaw-Boris,A., Yanagisawa
,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A.,
Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Kono,H., Nakamura
,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii
,Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,K., Shibata
,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander
,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.

```

```

TITLE    Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL  Nature 420, 563-573 (2002)
MEDLINE  22354683
PUBMED   12466851
COMMENT  Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Kawai,J., Kono,H., Miyazaki,A.,
T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Kono,H., Miyazaki,A.,
Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Sakai,K.,
Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami
,M., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y. Direct
Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)

```


RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. *Genome Res.* 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES	SOURCE
----------	--------

```

FEATURES
SOURCE
Location/Qualifiers
1. .443
/organism="Mus musculus"
/mol_type="mRNA"
/strain="ROD"
/db_xref="taxon:10090"
/clone="F830028K04"
/tissue_type="activated spleen"
/clone_lib="RIKEN full-length enriched, activated spleen"
97 a 128 c 117 g 101 t
BASE COUNT

```

Query Match	5.1%;	Score 50.8;	DB 13;	Length 443;
Best Local Similarity	62.7%;	Pred. No. 5.6e+02;		
Matches	79;	Conservative 0;	Mismatches 47;	Indels 0; Gaps 0

QY	597	CCTGCCAACATCTCCTTCCCTGCCAGCCAGACATCGGACTGGTTCTGTGTCAGGCTGA	656
Db			
	1	CCAGCCAACTTCTCCCTCCCCTGTCGCAGACCACCTGGTTGTTTTCAGTGCGAAGCTGAA	60
QY	657	AACAACGCCCAATGTCACAGCACAGCGCCCTCACAGTGGTGCCCCCAGGGTGTGACCAGAAG	716
Db	61	AACGATGTCGGTGTGGACAGAGTGCCCGCATCCCGCTGCCCGCCGAGAGCCCGAGCCCAAG	120
QY	717	ATGGAG	722
Db	121	CTGGTG	126

RESULT 42	CNS0091P	925 bp	DNA	linear	GSS 03-JUN-1999
LOCUS	CNS0091P				
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR19DJ16 of RCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL053013				
VERSION	AL053013.1	GI:4934461			
KEYWORDS	GSS.				
SOURCE	Drosophila melanogaster (fruit fly)				
ORGANISM	Drosophila melanogaster				
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 925)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)				
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osogawa and Aaron Mammosser in Pieter de Jong's laboratory in the Department of				

Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPci-98 and was constructed by partial EcoRI digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES	SOURCE
1. <i>General</i>	
2. <i>Specific</i>	
3. <i>Other</i>	
4. <i>Other</i>	
5. <i>Other</i>	
6. <i>Other</i>	
7. <i>Other</i>	
8. <i>Other</i>	
9. <i>Other</i>	
10. <i>Other</i>	
11. <i>Other</i>	
12. <i>Other</i>	
13. <i>Other</i>	
14. <i>Other</i>	
15. <i>Other</i>	
16. <i>Other</i>	
17. <i>Other</i>	
18. <i>Other</i>	
19. <i>Other</i>	
20. <i>Other</i>	
21. <i>Other</i>	
22. <i>Other</i>	
23. <i>Other</i>	
24. <i>Other</i>	
25. <i>Other</i>	
26. <i>Other</i>	
27. <i>Other</i>	
28. <i>Other</i>	
29. <i>Other</i>	
30. <i>Other</i>	
31. <i>Other</i>	
32. <i>Other</i>	
33. <i>Other</i>	
34. <i>Other</i>	
35. <i>Other</i>	
36. <i>Other</i>	
37. <i>Other</i>	
38. <i>Other</i>	
39. <i>Other</i>	
40. <i>Other</i>	
41. <i>Other</i>	
42. <i>Other</i>	
43. <i>Other</i>	
44. <i>Other</i>	
45. <i>Other</i>	
46. <i>Other</i>	
47. <i>Other</i>	
48. <i>Other</i>	
49. <i>Other</i>	
50. <i>Other</i>	
51. <i>Other</i>	
52. <i>Other</i>	
53. <i>Other</i>	
54. <i>Other</i>	
55. <i>Other</i>	
56. <i>Other</i>	
57. <i>Other</i>	
58. <i>Other</i>	
59. <i>Other</i>	
60. <i>Other</i>	
61. <i>Other</i>	
62. <i>Other</i>	
63. <i>Other</i>	
64. <i>Other</i>	
65. <i>Other</i>	
66. <i>Other</i>	
67. <i>Other</i>	
68. <i>Other</i>	
69. <i>Other</i>	
70. <i>Other</i>	
71. <i>Other</i>	
72. <i>Other</i>	
73. <i>Other</i>	
74. <i>Other</i>	
75. <i>Other</i>	
76. <i>Other</i>	
77. <i>Other</i>	
78. <i>Other</i>	
79. <i>Other</i>	
80. <i>Other</i>	
81. <i>Other</i>	
82. <i>Other</i>	
83. <i>Other</i>	
84. <i>Other</i>	
85. <i>Other</i>	
86. <i>Other</i>	
87. <i>Other</i>	
88. <i>Other</i>	
89. <i>Other</i>	
90. <i>Other</i>	
91. <i>Other</i>	
92. <i>Other</i>	
93. <i>Other</i>	
94. <i>Other</i>	
95. <i>Other</i>	
96. <i>Other</i>	
97. <i>Other</i>	
98. <i>Other</i>	
99. <i>Other</i>	
100. <i>Other</i>	

```

source
1. 925
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR19D16"
/clone_lib="xPCI-98"
/notes="end : TET3"
120 a 61 C 61 G 172 t 511 others
BASE COUNT
ORIGIN

```

[illegible]

RESULT 43	
CNS0052P	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	

JOURNAL

```
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D1070AG02NP1.
FEATURES
source
Location/Qualifiers
1..974
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1070YM03"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      59 a   235 c   80 g   310 t   280 others
ORIGIN
Query Match      4.8%; Score 47.2; DB 13; Length 974;
Best Local Similarity 25.8%; Pred. No. 1.3e+03;
Matches 85; Conservative 85; Mismatches 160; Indels 0; Gaps 0;
QY 663 GCCAATGCCAGCACACAGCCCTCACAGTGTGCCCCAGGTGTGTACCAAGATGAG 722
Db 909 GVVMMGGMGVGMVGMGGMMGGVGMGGMGGMGGMGVGMVGMVGMVGMVGMVGM 850
QY 723 GACTGGCAGAGTCCCTCGGAGAGCCCATCTTCCTTGCCTGCGCTCTACAGAGCAGCCGC 782
Db 849 GMRNGVAGMGMGVGMGGMGAGMGGMGGRGGMVGMVGMVGMVGMVGMVGMVGM 790
QY 783 CGTGTGAGTGAAGAGGAGTTTGGGGGTTTCAGGATAGGGAATGGGAGGTACAGAGCAGC 842
Db 789 MAGGGVGGNMAGMGGMVGMVGMVGMVGMVGMVGMVGMVGMVGMVGMVGMVGMV 730
QY 843 AAGCAGCAGCCATGTAGATGAACCTCCAGAGAGCAGCAGCAGCAGGAGTGCAGG 902
Db 729 GGMVGMVGMVGMVGMVGMVGMVGMVGMVGMVGMVGMVGMVGMVGMVGMVGMV 670
QY 903 CCATCAGCTGCACTGTCTGTATTTGAGTTTCATGCAAAATGAGTGTGTTTACGCTC 962
Db 669 VMCCCMVMMVMMVCCGGMVMMVMMVMMVMMVMMVMMVMMVMMVMMVMMVMMV 610
QY 963 TTGCCCAAAAAAAAAAAAAAAAAAAAAA 992
Db 609 ANAAAAAAAAAAAAAAAAAAAAAAAAA 580
RESULT 46
BZ772737/c
LOCUS
DEFINITION
mcs51f02.g1 HFOSMID007 Homo sapiens genomic, genomic survey
sequence.
ACCESSION
BZ772737
VERSION
BZ772737.1 GI:28946421
KEYWORDS
GSS.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 748)
Cook,L., Delehaunty,K., Fewell,G., Fulton,L., Magrini,V., Mardis,E.
, Miner,T., Nash,W., Williams,D. and Wilson,R.K.
Homo sapiens Fosmid End Reads
Unpublished
Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
```

```
Plate: mcv51 row: f column: 02
Class: fosmid ends
High quality sequence start: 36
High quality sequence stop: 550.
FEATURES
source
Location/Qualifiers
1..748
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone_lib="HFOSMID007"
/notes="Vector: pcc01fos; Site 1: Eco72i; Human whole
genome fosmid library was prepared at Washington
University Genome Sequencing Center. DNA was sheared for
blunt-ended ligation into pcc01fos inducible vector. DNA
was ordered from Coriell Cell Repository's DNA
polymorphism discovery resource."
BASE COUNT      211 a   133 c   104 g   299 t   1 others
ORIGIN
Query Match      4.7%; Score 46.4; DB 29; Length 748;
Best Local Similarity 58.1%; Pred. No. 1.8e+03;
Matches 100; Conservative 0; Mismatches 71; Indels 1; Gaps 1;
QY 813 AGCATAGGGAATGGGAGGTGACAGGAGCAAGCAGCAGCCATGTAGATGAACCTCC 872
Db 393 AGGAATGGGGAGATGTAGTTCGAGGGTACAAAATAGCAGATATGTAGATAAACAGGTC 334
QY 873 AGAGAGCAACGACGGCAGAGGACTGCAGGCCATCAGCGTGCACGTTCGTATTGGAGT 932
Db 333 TAGAATGTATGTATAACATGA-GGGAGATAACATAAACTTGCCTGTATTGCAAT 275
QY 933 TCATGCAAAATAGTGTGTTTGTAGTCTCTTGCACAAAAA 984
Db 274 TCATACAAATGAGTAGATTTTAACGTCTCTGCCACATGCAAAAAAGTAA 223
BX356664
BX356664 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1015YB03 3-PRIME, mRNA sequence.
ACCESSION
BX356664
VERSION
BX356664.1 GI:30378083
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D1015CA02NP1.
FEATURES
source
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1015YB03"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      116 a   88 c   93 g   398 t   506 others
```


[illegible]

BASE COUNT	257 a	170 c	162 g	96 t	250 others
ORIGIN	/note="end : 17"				

	Query Match	4.6%	Score 45.6;	DB 29;	Length 935;
	Best Local Similarity	25.8%	Pred. No. 2.1e+03;		
	Matches 109;	Conservative 105;	Mismatches 208;	Indels 0;	Gaps 0;
QY	342	GACCTGCTACCTACTCTCCGGGGCGTCTCCACCTCAGGTGCCCATGGACAGTGCC	401		
Db	513	SMCRTSSGSGCTCTCMYMSVSSCCSGCTCCGCGSSCTSKMCSCTYGCCKGCGGCS	572		
QY	402	AGGCTACAGATGCACCTGGGAGCTGTGTGTCAAAGCCAGTGTCTGAGCTGCGGGGCCAACTTC	461		
Db	573	TSCSSSSCCBBBSYSTCTBTCKSSGCTGTGGCTCCGGGGSGCGCGCGCGCGG	632		
QY	462	ACTCTGAGACAGAGGGGGCAGGCCCCACAGGGTGGAGATGATCTGCCAGGGCGTCTCTCGGGC	521		

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2004, 13:04:22 ; Search time 164 Seconds
(without alignments)
2669.830 Million cell updates/sec

Title: US-09-990-726-222

Perfect score: 992

Sequence:

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A COMB.seq.*
2: /cgn2_6/prodata/1/ina/5B COMB.seq.*
3: /cgn2_6/prodata/1/ina/6A COMB.seq.*
4: /cgn2_6/prodata/1/ina/6B COMB.seq.*
5: /cgn2_6/prodata/1/ina/6C COMB.seq.*
6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	992	100.0	992	4	US-09-996-243-222
2	47	4.7	87350	3	US-08-781-891-79
3	47	4.7	87350	4	US-09-618-166-79
4	47	4.7	87543	4	US-09-791-211-3
5	43.6	4.4	289	3	US-09-007-005-17
6	43.6	4.4	289	3	US-09-244-796-17
7	41.4	4.2	7218	1	US-08-232-463-14
8	40.6	4.1	480	4	US-09-252-991A-11640
9	40.6	4.1	1575	4	US-09-252-991A-11799
10	40.6	4.1	2022	4	US-09-252-991A-11716
11	40.2	4.1	1332	4	US-09-252-991A-10993
12	40.2	4.1	1656	4	US-09-252-991A-11258
13	40.2	4.1	2217	4	US-09-252-991A-11190
14	37.8	3.8	4411529	3	US-09-103-840A-2
15	37.8	3.8	1048	4	US-09-522-714-21
16	37.6	3.8	1396	1	US-08-123-161A-11
17	37.6	3.8	1396	1	US-08-483-278-11
18	37.6	3.8	1992	4	US-09-252-991A-5567
19	37.6	3.8	520	2	US-08-997-080-179
20	37.4	3.8	520	2	US-08-997-362-179
21	37.4	3.8	520	3	US-09-095-855-179
22	37.4	3.8	520	3	US-09-324-542-179
23	37.4	3.8	520	4	US-09-205-426-179
24	37.4	3.8	520	4	US-09-095-855-202
25	37.4	3.8	570	3	US-09-205-426-202
26	37.4	3.8	570	3	US-09-319-892-3
27	36.6	3.7	1273	3	Sequence 3, Appli

28	36.6	3.7	1273	4	US-09-727-238-3	Sequence 3, Appli
29	36.4	3.7	801	4	US-09-252-991A-11102	Sequence 1102, A
30	36.4	3.7	1116	4	US-09-252-991A-11158	Sequence 11158, A
31	36.4	3.7	1584	4	US-09-252-991A-11094	Sequence 11094, A
32	36.4	3.7	2028	2	US-09-211-930-12	Sequence 12, Appl
33	36.4	3.7	2028	3	US-09-340-993-12	Sequence 12, Appl
34	36.4	3.7	2028	4	US-09-468-442-12	Sequence 12, Appl
35	36.2	3.6	570	4	US-09-252-991A-5589	Sequence 5589, Ap
36	36.2	3.6	1545	4	US-09-252-991A-5553	Sequence 5553, Ap
37	36.2	3.6	1914	4	US-09-252-991A-5670	Sequence 5670, Ap
38	36	3.6	2625	3	US-09-245-041-18	Sequence 18, Appl
39	36	3.6	4403765	3	US-09-103-840A-2	Sequence 2, Appli
40	36	3.6	4411529	3	US-09-103-840A-1	Sequence 1, Appli
41	35.8	3.6	28958	1	US-08-258-261B-6	Sequence 6, Appli
42	35.8	3.6	28958	1	US-08-456-837-6	Sequence 6, Appli
43	35.8	3.6	28958	1	US-08-457-342-6	Sequence 6, Appli
44	35.8	3.6	28958	1	US-08-457-646A-6	Sequence 6, Appli
45	35.8	3.6	28958	1	US-08-458-076A-6	Sequence 6, Appli
46	35.8	3.6	28958	1	US-08-764-233A-4	Sequence 4, Appli
47	35.8	3.6	28958	1	US-08-457-335A-6	Sequence 6, Appli
48	35.8	3.6	28958	3	US-09-028-934-6	Sequence 6, Appli
49	35.8	3.6	28958	1	US-08-729-214-6	Sequence 6, Appli
50	35.8	3.6	49377	1	US-08-764-233A-1	Sequence 1, Appli
51	35.6	3.6	366	4	US-09-775-932-13	Sequence 13, Appl
52	35.6	3.6	387	4	US-09-775-932-11	Sequence 11, Appl
53	35.6	3.6	581	2	US-08-461-030C-1	Sequence 1, Appli
54	35.6	3.6	588	3	US-08-744-138-1	Sequence 1, Appli
55	35.6	3.6	588	4	US-09-241-376-1	Sequence 1, Appli
56	35.4	3.6	294	4	US-09-252-991A-10962	Sequence 10962, A
57	35.4	3.6	447	4	US-09-252-991A-8928	Sequence 8928, Ap
58	35.4	3.6	676	2	US-08-533-669A-3	Sequence 3, Appli
59	35.4	3.6	676	4	US-09-183-861-3	Sequence 3, Appli
60	35.4	3.6	676	4	US-09-022-765-3	Sequence 3, Appli
61	35.4	3.6	676	4	US-09-551-974A-3	Sequence 3, Appli
62	35.4	3.6	696	4	US-09-252-991A-9236	Sequence 9236, Ap
63	35.4	3.6	1011	4	US-09-252-991A-8814	Sequence 8814, Ap
64	35.4	3.6	1125	4	US-09-252-991A-14001	Sequence 14001, A
65	35.4	3.6	1617	4	US-09-252-991A-14135	Sequence 14135, A
66	35.4	3.6	1728	4	US-09-252-991A-9152	Sequence 9152, Ap
67	35.2	3.5	23673	4	US-09-773-816-1	Sequence 1, Appli
68	35	3.5	588	4	US-09-205-258-64	Sequence 64, Appl
69	35	3.5	1683	3	US-09-347-803-11	Sequence 11, Appl
70	34.8	3.5	1646	4	US-09-443-041A-19	Sequence 19, Appl
71	34.8	3.5	1869	1	US-07-937-608-28	Sequence 28, Appl
72	34.8	3.5	1869	3	US-08-029-170-28	Sequence 28, Appl
73	34.6	3.5	816	4	US-09-252-991A-11434	Sequence 11434, A
74	34.6	3.5	1116	4	US-09-522-714-31	Sequence 31, Appl
75	34.6	3.5	1230	4	US-09-461-325-80	Sequence 80, Appl
76	34.6	3.5	1254	4	US-09-252-991A-11478	Sequence 11478, A
77	34.6	3.5	1254	4	US-09-252-991A-11455	Sequence 11455, A
78	34.6	3.5	2028	4	US-09-252-991A-11416	Sequence 11416, A
79	34.6	3.5	2346	3	US-09-193-503B-5	Sequence 5, Appli
80	34.6	3.5	2346	4	US-09-415-839-5	Sequence 5, Appli
81	34.6	3.5	3471	5	US-09-415-839-5	Sequence 5, Appli
82	34.6	3.5	3530	3	US-08-704-711A-10	Sequence 2, Appli
83	34.6	3.5	3530	4	US-09-521-220-10	Sequence 10, Appl
84	34.6	3.5	4190	3	US-08-538-291A-2	Sequence 2, Appli
85	34.6	3.5	4190	4	US-09-589-619-2	Sequence 2, Appli
86	34.2	3.4	423	1	US-08-470-179-186	Sequence 186, App
87	34.2	3.4	1863	4	US-09-252-991A-3225	Sequence 3225, Ap
88	34.2	3.4	2154	4	US-09-252-991A-3346	Sequence 3346, Ap
89	34.2	3.4	9208	4	US-09-068-506-1	Sequence 1, Appli
90	34	3.4	2461	1	US-08-282-141-1	Sequence 1, Appli
91	34	3.4	3402	4	US-09-252-991A-1374	Sequence 1374, Ap
92	34	3.4	3687	4	US-09-252-991A-1193	Sequence 1193, Ap
93	34	3.4	4266	4	US-09-252-991A-1234	Sequence 1234, Ap
94	33.8	3.4	1518	3	US-09-240-915-2	Sequence 2, Appli
95	33.8	3.4	1518	3	US-09-591-435-2	Sequence 2, Appli
96	33.8	3.4	1536	4	US-09-252-991A-11652	Sequence 11652, A
97	33.8	3.4	1656	4	US-09-228-986-63	Sequence 63, Appl
98	33.8	3.4	1761	4	US-09-252-991A-11704	Sequence 11704, A
99	33.8	3.4	1767	4	US-09-252-991A-7614	Sequence 7614, Ap
100	33.8	3.4	2118	4	US-09-252-991A-11782	Sequence 11782, A

c 101 33.8 3.4 3447 2 US-08-313-185-57
c 102 33.8 3.4 3447 3 US-09-082-614A-57
103 33.6 3.4 2572 3 US-08-804-439A-9
104 33.6 3.4 2572 3 US-08-720-229-9
c 105 33.6 3.4 2567 4 US-09-185-501B-12
106 33.6 3.4 15652 4 US-09-422-936-60
107 33.4 3.4 407 4 US-08-948-113D-11
108 33.4 3.4 1053 4 US-09-252-991A-14223
c 109 33.4 3.4 1846 1 US-08-483-389-117
c 110 33.4 3.4 10660 2 US-08-267-803B-8
c 111 33.4 3.4 10660 3 US-09-041-886-16
c 112 33.2 3.3 465 4 US-09-252-991A-12679
113 33.2 3.3 750 4 US-09-252-991A-13279
c 114 33.2 3.3 858 4 US-09-252-991A-8589
c 115 33.2 3.3 1897 1 US-08-184-632-1
116 33.2 3.3 1947 1 US-07-951-715A-3
117 33.2 3.3 1947 2 US-08-459-448A-3
118 33.2 3.3 1947 3 US-08-459-595A-3
119 33.2 3.3 1947 3 US-08-459-595A-3
120 33.2 3.3 1947 3 US-08-459-595A-3
121 33.2 3.3 1947 3 US-08-459-595A-3
122 33.2 3.3 1947 3 US-08-459-595A-3
123 33.2 3.3 2190 2 US-08-492-027A-7
124 33.2 3.3 2194 2 US-08-948-569A-9
125 33.2 3.3 2194 3 US-09-188-469-9
c 126 33.2 3.3 2194 3 US-09-397-238A-9
127 33.2 3.3 2502 4 US-09-252-991A-8656
128 33.2 3.3 2852 4 US-09-027-137-2
129 33.2 3.3 2852 4 US-09-344-441-2
130 33.2 3.3 3468 1 US-07-951-715A-2
131 33.2 3.3 3468 1 US-07-951-715A-4
132 33.2 3.3 3468 1 US-07-951-715A-8
133 33.2 3.3 3468 2 US-08-459-448A-2
134 33.2 3.3 3468 2 US-08-459-448A-4
135 33.2 3.3 3468 2 US-08-459-448A-8
136 33.2 3.3 3468 3 US-08-459-595A-2
137 33.2 3.3 3468 3 US-08-459-595A-4
138 33.2 3.3 3468 3 US-08-459-595A-8
139 33.2 3.3 3468 3 US-08-459-595A-8
140 33.2 3.3 3468 3 US-08-459-595A-8
141 33.2 3.3 3468 3 US-08-459-595A-8
142 33.2 3.3 3468 3 US-08-459-595A-8
143 33.2 3.3 3468 3 US-08-459-595A-8
144 33.2 3.3 3468 3 US-08-459-595A-8
145 33.2 3.3 3468 3 US-09-053-549-3
146 33.2 3.3 3468 3 US-09-053-549-5
147 33.2 3.3 3468 4 US-09-547-422-2
148 33.2 3.3 3468 4 US-09-547-422-4
149 33.2 3.3 3468 4 US-09-547-422-8
150 33.2 3.3 3546 1 US-07-951-715A-10

ALIGNMENTS

RESULT 1
US-09-996-243-222
; Sequence 222, Application US/09996243
; Patent No. 647825
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.

; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C13
; CURRENT APPLICATION NUMBER: US/09/996,243
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738

;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088742
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088810
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088858
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089440
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089600
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089908
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089948
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089952
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090246
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090252
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090254
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090355
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24

;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Query Match 100.0%; Score 992; DB 4; Length 992;

Best Local Similarity 100.0%; Pred. No. 1.4e-243;

Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACGAGCCAGAACTAGGAGTTCTCACTCCCGAGCAGAGGCCCTACACCCACCGAG 60
Db 1 GGCACGAGCCAGAACTAGGAGTTCTCACTCCCGAGCAGAGGCCCTACACCCACCGAG 60
QY 61 GCATGGGGCTCCCTGGGCTGTTCTGCTGGCGTCTGCTGCCAGCAGCTTCTCCAGG 120
Db 61 GCATGGGGCTCCCTGGGCTGTTCTGCTGGCGTCTGCTGCCAGCAGCTTCTCCAGG 120
QY 121 CACGGGAGGAGAAATTACCCCTGTTCTCAATTCCTTACAAAGTCTGGAAGTTTCC 180
Db 121 CACGGGAGGAGAAATTACCCCTGTTCTCAATTCCTTACAAAGTCTGGAAGTTTCC 180
QY 181 CCAAGGCGGTGGGTGCTCATACCTGTGTGCACCCCGAGCCACCGCCCATCACT 240
Db 181 CCAAGGCGGTGGGTGCTCATACCTGTGTGCACCCCGAGCCACCGCCCATCACT 240
QY 241 ATTCCCTCTGTGGAAACCAAGAACATCAAGGTGGGCCAAGAGTGTGAAGACCCACGAGC 300
Db 241 ATTCCCTCTGTGGAAACCAAGAACATCAAGGTGGGCCAAGAGTGTGAAGACCCACGAGC 300
QY 301 CGGCGCTCTTCAACTCAACGTCACTCAAGTCCAGTCCAGCTGCTCACTACTTCT 360
Db 301 CGGCGCTCTTCAACTCAACGTCACTCAAGTCCAGTCCAGCTGCTCACTACTTCT 360
QY 361 GCGGGCGTCTCCACCTCAGGTGCCATGTGGAGCAGTCCAGCTGCTCACTACTTCT 420
Db 361 GCGGGCGTCTCCACCTCAGGTGCCATGTGGAGCAGTCCAGCTGCTCACTACTTCT 420
QY 421 AGCTGTGTCCAAGCCAGTGTCTGAGTGGCGGCAAACTTCACTGTGAGACAGAGGG 480

Db 421 AGCTGTGTCAGCCAGTGTCTGAGTGTGCGGCGCAACTTCACTCTGCAGGACAGAGGG 480
QY 481 CAGGCCCCAGGGTGGAGATGATCTGCCAGGCGTCTCGGGCAGCCACCTATCAACA 540
Db 481 CAGGCCCCAGGGTGGAGATGATCTGCCAGGCGTCTCGGGCAGCCACCTATCAACA 540
QY 541 GCCTGATCGGGAAGATGGGCGAGTCCCTCAGCAGAGACCATGCCACAGGCGAGCTG 600
Db 541 GCCTGATCGGGAAGATGGGCGAGTCCCTCAGCAGAGACCATGCCACAGGCGAGCTG 600
QY 601 CCAACTTCTCTCTCTCGGCGCAGACATCGGACTGTTCTGTGCCAGGCTGCAACA 660
Db 601 CCAACTTCTCTCTCTCGGCGCAGACATCGGACTGTTCTGTGCCAGGCTGCAACA 660
QY 661 AGCCCAATGTCCAGCACAGCGCCCTCACAGTGTGTGCCCGCCAGGTGGTACCAGAGATGG 720
Db 661 AGCCCAATGTCCAGCACAGCGCCCTCACAGTGTGTGCCCGCCAGGTGGTACCAGAGATGG 720
QY 721 AGGACTGGCAGGTCCTCTGGAGAGCCCATCTTGGCTTGCCTCTACAGGACACCC 780
Db 721 AGGACTGGCAGGTCCTCTGGAGAGCCCATCTTGGCTTGCCTCTACAGGACACCC 780
QY 781 GCGCTCTGAGTCAAGCAGGATTTGGGGGTTTCAGGATAGGGAATGGGGAGTTCAGGAC 840
Db 781 GCGCTCTGAGTCAAGCAGGATTTGGGGGTTTCAGGATAGGGAATGGGGAGTTCAGGAC 840
QY 841 GCAAGCAGCGCCATGTAGTAATGAACCGTCCAGAGAGCCCAAGCAGCGGAGGACTGCA 900
Db 841 GCAAGCAGCGCCATGTAGTAATGAACCGTCCAGAGAGCCCAAGCAGCGGAGGACTGCA 900
QY 901 GGCATCAGCGTGCAGTCTCGTATTTGGAGTTCATGCAAAATGAGTGTGTTAGTGC 960
Db 901 GGCATCAGCGTGCAGTCTCGTATTTGGAGTTCATGCAAAATGAGTGTGTTAGTGC 960
QY 961 TCTTGGCCACAAAAA992
Db 961 TCTTGGCCACAAAAA992

RESULT 2

US-08-781-891-79/c
; Sequence 79, Application US/08781891
; Patent No. 6090620
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; APPLICANT: Yu, Chang-En
; APPLICANT: Oshima, Junko
; APPLICANT: Mulligan, John T.
; APPLICANT: Schellenberg, Gerald D.
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
; TITLE OF INVENTION: WERNER'S SYNDROME
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781.891
; FILING DATE: 27-DEC-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: 6090620tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 240052.419
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 79:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 87350 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-781-891-79

Query Match 4.7%; Score 47; DB 3; Length 87350;
Best Local Similarity 78.9%; Pred. No. 0.045;
Matches 56; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 922 GTATTGGAGTTTCATGCAAAATGAGTGTGTTTGTAGTCTTTCGCCACAAAAA981
Db 18894 GTATTGGAGTTTCATGCAAAATGAGTGTGTTTGTAGTCTTTCGCCACAAAAA981
QY 982 AAAAAAAAAA992
Db 18834 ATGTTAAACAA 18824

RESULT 3

US-09-618-166-79/c
; Sequence 79, Application US/09618166
; Patent No. 6583112
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; APPLICANT: Yu, Chang-En
; APPLICANT: Oshima, Junko
; APPLICANT: Mulligan, John T.
; APPLICANT: Schellenberg, Gerald D.
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
; TITLE OF INVENTION: WERNER'S SYNDROME
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed Intellectual Property Law Group
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/618.166
; FILING DATE: 17-Jul-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Mcmasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 240052.419C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 79:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 87350 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 79:
US-09-618-166-79

Query Match 4.7%; Score 47; DB 4; Length 87350;
Best Local Similarity 78.9%; Pred. No. 0.045;
Matches 56; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 922 GTATTGGAGTTTCATGCAAAATGAGTGTGTTTGTAGTCTTTCGCCACAAAAA981

Db 18894 GTATTGGATTCCCTGCTAAATGAGTGGATTTCAGCTGCTTTGCCACAAAACAAAAA 18835
QY 982 AAAAAAAAAA 992
Db 18834 ATGGTTAACAA 18824

RESULT 4

US-09-791-211-3/C
Sequence 3, Application US/09791211
Patent No. 6448080
GENERAL INFORMATION:
APPLICANT: Donna T. Ward
APPLICANT: Andrew T. Watt
TITLE OF INVENTION: ANTISENSE MODULATION OF WEN EXPRESSION
FILE REFERENCE: RTS-0205
CURRENT APPLICATION NUMBER: US/09/791,211
CURRENT FILING DATE: 2001-02-23
NUMBER OF SEQ ID NOS: 90
SEQ ID NO 3
LENGTH: 87543
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: unsure
LOCATION: 7421
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 7427
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 11609
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 12605
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 12742
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 29370
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 29422
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 29979
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 29980
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 29981
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 30136
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 30140
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 31205
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 31206
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 31592
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 33095
OTHER INFORMATION: unknown

NAME/KEY: unsure
LOCATION: 33160
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 34066
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 34072
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 36816
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 39020
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 42164
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 42459
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 46808
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 46823
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 46826
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 47291
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 52786
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 53384
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 54684
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 59215
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 59235
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 59242
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 63290
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 66614
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 68660
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 68697
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 68718
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 68733
OTHER INFORMATION: unknown
NAME/KEY: unsure

```
/ LOCATION: 68739
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 69785
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 79134
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 79198
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 8636
/ OTHER INFORMATION: unknown
/ US-09-791-211-3

Query Match
Best Local Similarity 4.7%; Score 47; DB 4; Length 87543;
Matches 56; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 922 GTATTGGAGTTCATGCAAAATGAGTGTGTTTAGCTGCTCTTGCCACAAAAA 981
Db 19087 GTATTGGAGTTCGCTTAATGAGTGGATTCAGCTGCTTTGCCACAAAA 19028

QY 982 AAAAAA 992
Db 19027 ATGGTTACAA 19017

RESULT 5
US-09-007-005-17
; Sequence 17, Application US/09007005B
; Patent No. 6258558
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihe
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; FILE REFERENCE: 00786/350003
; CURRENT APPLICATION NUMBER: US/09/007,005B
; CURRENT FILING DATE: 1998-01-14
; EARLIER FILING DATE: 1997-01-27
; EARLIER FILING DATE: 1997-01-27
; EARLIER FILING DATE: 1997-11-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; NAME/KEY: misc_feature
; LOCATION: (1)...(289)
; OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17

Query Match
Best Local Similarity 4.4%; Score 43.6; DB 3; Length 289;
Matches 27; Conservative 85; Mismatches 111; Indels 0; Gaps 0;

QY 769 ACAGGACGCCCGCTGAGTGAAGAGGAGTTGGGGGTTTCAGATAGGATCGGG 828
Db 65 RNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRN 124
QY 829 AGGTGAGGACGCAAGACGACGACCATGTAGATGAACCGTCAGAGACCCAGCACGG 888
Db 125 RNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRN 184
QY 889 CAGGAGCTCAGCGCCATCAGCGTCTGCTTATTTGGAGTTTCATGAAATGAGTG 948
Db 185 RNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRN 244
QY 949 TGTTTAGCTGCTTGGCCACAAAAA 991
Db 245 RCRURGRGRCRUAAAAA 287

RESULT 6
US-09-244-796-17
; Sequence 17, Application US/09244796
; Patent No. 6281344
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihe
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; FILE REFERENCE: 00786/350007
; CURRENT APPLICATION NUMBER: US/09/244,796
; CURRENT FILING DATE: 1999-02-05
; EARLIER FILING DATE: 1997-01-27
; EARLIER FILING DATE: 1997-01-27
; EARLIER FILING DATE: 1997-11-06
; EARLIER FILING DATE: 1997-11-06
; EARLIER FILING DATE: 1998-01-14
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; NAME/KEY: misc_feature
; LOCATION: (1)...(289)
; OTHER INFORMATION: n = A,T,C or G
US-09-244-796-17

Query Match
Best Local Similarity 4.4%; Score 43.6; DB 3; Length 289;
Matches 27; Conservative 85; Mismatches 111; Indels 0; Gaps 0;

QY 769 ACAGGACGCCCGCTGAGTGAAGAGGAGTTGGGGGTTTCAGATAGGATCGGG 828
Db 65 RNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRN 124
QY 829 AGGTGAGGACGCAAGACGACGACCATGTAGATGAACCGTCAGAGACCCAGCACGG 888
Db 125 RNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRN 184
QY 889 CAGGAGCTCAGCGCCATCAGCGTCTGCTTATTTGGAGTTTCATGAAATGAGTG 948
Db 185 RNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRN 244
QY 949 TGTTTAGCTGCTTGGCCACAAAAA 991
Db 245 RCRURGRGRCRUAAAAA 287

RESULT 7
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
```

```
QY 889 CAGAGGACTGCAGGCCATCAGGTGCACCTGTTCTGATTGGAGTTTCATGCAAAATGAGTG 948
Db 185 RNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRN 244
QY 949 TGTTTAGCTGCTTGGCCACAAAAA 991
Db 245 RCRURGRGRCRUAAAAA 287

RESULT 6
US-09-244-796-17
; Sequence 17, Application US/09244796
; Patent No. 6281344
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihe
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; FILE REFERENCE: 00786/350007
; CURRENT APPLICATION NUMBER: US/09/244,796
; CURRENT FILING DATE: 1999-02-05
; EARLIER FILING DATE: 1997-01-27
; EARLIER FILING DATE: 1997-01-27
; EARLIER FILING DATE: 1997-11-06
; EARLIER FILING DATE: 1997-11-06
; EARLIER FILING DATE: 1998-01-14
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; NAME/KEY: misc_feature
; LOCATION: (1)...(289)
; OTHER INFORMATION: n = A,T,C or G
US-09-244-796-17

Query Match
Best Local Similarity 4.4%; Score 43.6; DB 3; Length 289;
Matches 27; Conservative 85; Mismatches 111; Indels 0; Gaps 0;

QY 769 ACAGGACGCCCGCTGAGTGAAGAGGAGTTGGGGGTTTCAGATAGGATCGGG 828
Db 65 RNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRN 124
QY 829 AGGTGAGGACGCAAGACGACGACCATGTAGATGAACCGTCAGAGACCCAGCACGG 888
Db 125 RNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRN 184
QY 889 CAGGAGCTCAGCGCCATCAGCGTCTGCTTATTTGGAGTTTCATGAAATGAGTG 948
Db 185 RNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRN 244
QY 949 TGTTTAGCTGCTTGGCCACAAAAA 991
Db 245 RCRURGRGRCRUAAAAA 287

RESULT 7
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
```


Best Local Similarity		44.4%;	Pred. NO. 0.44;	
Matches 163;		Conservative	0;	Mismatches 204; Indels 0; Gaps 0;
QY	260	GAACATCAAGGTGGCCCAAGAGGTGGTGAAGACCCAGAGCGGGCTCTCTCAACCTCAA	319	
Db	1173	GGAGATCAAGAAGACCATCGAGCAGGTGGTGCAGCGGGCTGTACTACTTGCCTCGG	1232	
QY	320	CGTCACACTCAAGTCCAGTCCAGACCTGCTCACTACTTCTGCGGGCTCTCCACATC	373	
Db	1233	CGTTCCGGACCTGCACATCCGCACTCGACAGTACTCTTCCACCTATTTCGCGGGCTC	1292	
QY	380	AGTGGCCCATGTGCACAGTGCAGGCTACAGATGCATGGGAGCTGTGGTCCAAAGCCAGT	439	
Db	1293	CGGCGGCATGGGCCACCCGGAGCGGATCAAGATCCTCAAGTGTCTTGGGAGCGGCATCG	1352	
QY	440	GTCTGAGCTGGGGCCAACTTCACTTGCAGGAAGAGGGGSCAGGCCCCAGGGTGGAGAT	499	
Db	1353	CAGCGAGTTTCGGCGGGCCGCCACGAGCTGTACGAGATCAACTACGCCGGCAGCAGGACGA	1412	
QY	500	GATCTGCCAGCGTCCTCGGGCAGCCACACTATCAACCAAGCCTGATCGGGAAGGATGG	559	
Db	1413	GATCCGCATGCAAGCCTTCGGCCAGGCGGTGGCAGCGGGCGGATGAAGGCGCATGCTCGG	1472	
QY	560	GCAGGTCACACTGCAGCAGAGACCATGCCACAGGCAGCCTGCCACATCTTCTTCTCTGCC	619	
Db	1473	CATGGTCGACAGTGAATGGCGGACTACGAGAGACGGCTGGACCGTCGCGACCTGCA	1532	
QY	620	GAGCCAG	626	
Db	1533	CAACCCG	1539	

```

RESULT 10
US-09-252-991A-11716
; Sequence 11716, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.1136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11716
; LENGTH: 2022
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11716

```

Query Match	4.1%	Score 40.6;	DB 4;	Length 2022;
Best Local Similarity	44.4%;	Pred. No. 0.48;		
Matches 163;	Conservative 0;	Mismatches 204;	Indels 0;	Gaps 0;
QY	260	GAACATCAAGTGGCCAAAGAAGTGGTGAAGACCCACAGACGGCGCTCCTTCAACCTCAA	319	
Db	1078	GGAGATCAAGAAGACATCGAGCAGTGGTGCACAGCGGCTGATCTACCTGGCCTCCGG	1137	
QY	320	CGTCACACTCAAGTCCAGTCCAGACCTGGTCACTCTTTCGGCGGCGCTCTCCACCTC	379	
Db	1138	CGTTCGGACTCGACAAATCGCAACTCGACAGTACCTCTCCACCTATTGCCGCGGCTC	1197	
QY	380	AGGTGCCCATGTGGACAGTGCACAGGTACAGATGCACCTGGGAGCTGTGTCTCAAGCCAGT	439	
Db	1198	CGCGCGCATGGGCCACCGGGAGCGGATCAAGATCCTCAAGTCTCTGGAGCGCCATCGG	1257	
QY	440	GTCTGAGCTCGGGCCACATTCTCTGTGAGGACAGAGGGGCGAGGCCCGAGGTGGAGAT	499	
Db	1258	CACCGAGTTTCGGCGGGCCGCCACGAGTGTACAGATCACTACGCGCGGACGACGACGA	1317	

Qy	500	GATCTGCCAGCGTCTCGGGCGGCCACCTATCACCAACAGCCTGTATCGGGAAGATGG	559
Db	1318	GATCCGCATGACGGCCTCGCCACAGGCGGTGGGAGCGGGCGATGAAGGCGCATGCTCGG	1377
Qy	560	GCAGGTCCACCTGCAGCAGAGACATCCACAGCGACGCTGCCAACATTCTCTTCCTGCC	619
Db	1378	CATGGTCGACAGTGCATGGCGCACTACGACGAGACGGCTGGACCTGCGCACCTGCA	1437
Qy	620	GAGCCAG	626
Db	1438	CAACCCG	1444

```

RESULT 11
US-09-252-991A-10993
; Sequence 10993, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenstein et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMIN
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNO
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 10993
; LENGTH: 1332
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-10993

```

Query Match	4.1%;	Score 40.2;	DB 4;	Length 1332;
Best Local Similarity	49.8%;	Prod. No. 0.53;		
Matches 102;	Conservative 0;	Mismatches 103;	Indels 0;	Gaps 0;
QY	218	CCAGCCACCACGCCCATCACCTATTCCCTCTGGAACCAAGACATCAAGTGGCAA	277	
Db	826			
QY	278	GAAGGTGGTGAAGACCCACGAGCGCGCTCTTCAACCTCAAGCTCACATCAAGTCCAG	337	
Db	886	GCGGATGGCGATGGTCGCGTGACCGCCACCATCAGGTCTAGCGCAACATGACCACTCC	945	
QY	338	TCCAGACCTGCTCACCTACTTCTGCGGGCGTCTCCACTCAGGTGCCATGTGACAG	397	
Db	946	GCGGATCAGCGGGAGAGGCGCTGGTAGCGCTCTCCAGCACTGGTCACGGTGAACAC	1005	
QY	398	TGCCAGGCTACAGATGCACTGGGAG	422	
Db	1006	CGCGACGCGCGGATCTGCGCGAG	1030	

RESULT 12
US-09-252-991A-11258/c
; Sequence 11258, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142

```
; SEQ ID NO 11258
; LENGTH: 1656
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11258

Query Match
Best Local Similarity 4.1%; Score 40.2; DB 4; Length 1656;
Matches 102; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 218 CCAGCCACCCGCGCCATCCTTCCCTCTGTGGAAACCAAGACATCAAGTGGCCAA 277
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
570 CCGCCACCCGAGAGATCGCGGGATCCCGGGATCCGATACCACTCGGTGACCGA 511
QY 278 GAAGTGGTGAAGACCCACGAGCGGCTCTTCAACCTCAAGTCACTCAAGTCCAG 337
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
510 GCCGATGCCATGGTTCGGGTGACCGCCACCATCAGTCTAGCGACCATGACCCCC 451
QY 338 TCAGACCTGCTCACCCTACTTCTGCGGGGCTCTCCACCTCAGGTGCCCATGTGGACAG 397
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
450 GCCGATCAGCGCGAGGAGGCTGGTAGCGCTGCTCCAGCACCTGGGTACGGTGAACAC 391
QY 398 TGCAGGCTACATGCATGGGAG 422
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
390 CCGCAGGCGCGCATGTGCCGCGAG 366

RESULT 13
US-09-252-991A-11190/c
; Sequence 11190, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11190
; LENGTH: 2217
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (2217)
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-252-991A-11190

Query Match
Best Local Similarity 4.1%; Score 40.2; DB 4; Length 2217;
Matches 102; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 218 CCAGCCACCCGCGCCATCCTTCCCTCTGTGGAAACCAAGACATCAAGTGGCCAA 277
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
778 CCGCCACCCGAGAGATCGCGGGATCCCGGGATCCGATACCACTCGGTGACCGA 719
QY 278 GAAGTGGTGAAGACCCACGAGCGGCTCTTCAACCTCAAGTCACTCAAGTCCAG 337
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
718 GCCGATGCCATGGTTCGGGTGACCGCCACCATCAGTCTAGCGACCATGACCCCC 659
QY 338 TCAGACCTGCTCACCCTACTTCTGCGGGGCTCTCCACCTCAGTGGCCCATGTGGACAG 397
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
658 GCCGATCAGGCGCGAGGAGGCTGGTAGCGCTGCTCCAGCACCTGGGTACGGTGAACAC 599
QY 398 TGCAGGCTACATGCATGGGAG 422
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
598 CCGCAGGCGCGCATGTGCCGCGAG 574
```

```
RESULT 14
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match
Best Local Similarity 3.8%; Score 37.8; DB 3; Length 4403765;
Matches 114; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

QY 338 TCAGACCTGCTCACCCTACTTCTGCGGGGCTCTCCACCTCAGGTGCCCATGTGGACAG 397
Db 284113 TTCCGACCTGTGACATCTCTCGGTTCGGGTACGCGGCGCGGATCCGGCCAC 284054
QY 398 TGCCAGGCTACAGATGCACTGGGAGCTGTGTTCCAGCCAGTGTGAGCTGGGCGCAA 457
Db 284053 CGCTTGGACCGCGCGAGCGGTGTGCAACAAACCCCGCCGACCTGACGTGAC 283994
QY 458 CTTCACTTTCAGGACAGAGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 517
Db 283993 CCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 283934
QY 518 GGGCAGCGCCACCTATCACCACAGCTGATCGGGAAGATGGGCAAGTCCACCTGAGCA 577
Db 283933 GGCCATCCGAGCGGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 283874
QY 578 G 578
Db 283873 G 283873

RESULT 15
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
```

323 CACACTCAAGTCCAGTCCAGACCTGCTACCTACTTCTGCGGGCGTCTCCACCTCAGG 382

Db 1309 GGGTGGGTCAGACTGTCTGGAGTAAGACATTTCAGATAAATATCTGCTCTGCTCA 1368
QY 969 CAAAAAATAAAAAAAAAAAAAA 992
Db 1369 CAAAAAATAAAAAAAAAAAAAA 1392

RESULT 18
US-08-483-278-11
; Sequence 11, Application US/08483278
; Patent No. 5686073
; GENERAL INFORMATION:
; APPLICANT: Campbell, Kevin P.
; APPLICANT: Ibrahimov, Oxana B.
; APPLICANT: Ervasti, James M.
; APPLICANT: Leveille, Cynthia J.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING DYSTROPHIN-ASSOCIATED
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kevin M. Farrell, P.C.
; STREET: P.O. Box 999
; CITY: York Harbor
; STATE: ME
; COUNTRY: USA
; ZIP: 03911
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: System In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,278
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/123,161
; FILING DATE: 16-SEP-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Farrell, Kevin M.
; REGISTRATION NUMBER: 35,505
; REFERENCE/DOCKET NUMBER: UIRP89-11A5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (207) 363-0558
; TELEFAX: (207) 363-0528
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1396 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4...1164

US-08-483-278-11
Query Match 3.8%; Score 37.6; DB 1; Length 1396;
Best Local Similarity 65.5%; Pred. No. 2.5;
Matches 55; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
QY 909 GCGTGCACTGTTTCGTATTTGGAGTTCATGCAAAATGAGTGTCTTTTAGCTGCTCTTGCCA 968
Db 1309 GGGTGGGTCAGACTGTCTGGAGTAAGACATTTCAGATAAATATCTGCTCTGCTCA 1368
QY 969 CAAAAAATAAAAAAAAAAAAAA 992
Db 1369 CAAAAAATAAAAAAAAAAAAAA 1392

RESULT 19
US-09-252-991A-5567/c
; Sequence 5567, Application US/09252991A

; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 5567
; LENGTH: 1992
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-5567

Query Match 3.8%; Score 37.6; DB 4; Length 1992;
Best Local Similarity 50.6%; Pred. No. 2.8;
Matches 91; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 515 CTCGGGCAGCCACCTATCACCAACAGCCTGATCGGGAAGGATGGCAGGTCACCTGCA 574
Db 1569 CTCGTGCGCGCGCGCGCTCCATGAACAGAGATCAGGGCTGCGGTAGAGCGGTAGAA 1510
QY 575 GCAGAGACCATGCCACAGCAGCCTGCCAACTTCTCTTCTCCGCGAGCCAGACATCGGA 634
Db 1509 GAGCAGTCGCGGACCAAGCTCCAGCGGCCCATCTCTCGGGGCCGACGGCTTCGTCAT 1450
QY 635 CTGGTTCTGTCGTCAGGGCTGCAAAACGCCCAATGTCAGACACAGCCCTCACAGTGGT 694
Db 1449 CTGCTGCTGAGCAGGCTTTCAGGAATATCCAGGCGCAGGCGGACAGGGTGGT 1390

RESULT 20
US-08-997-080-179/c
; Sequence 179, Application US/08997080
; Patent No. 5968524
; GENERAL INFORMATION:
; APPLICANT: WATSON, JAMES D.
; APPLICANT: TAN, PAUL L.J.
; TITLE OF INVENTION: METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY-
; NUMBER OF SEQUENCES: 194
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/997,080
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 179:

SEQUENCE CHARACTERISTICS:
LENGTH: 520 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-997-080-179

Query Match
Best Local Similarity 3.8%; Score 37.4; DB 2; Length 520;
Matches 107; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 374 CACCTCAGTGCCTCCATGTGACAGTCCAGGCTACAGATGCACTGGGAGCTGTGTCCTCAA 433
DB 324 CTCCTCGGTGAGCTCGGGATCGACACCGGATGATGTCCTGCTGGTGGATTGAC 265
QY 434 GCCAGTGTCTGAGCTGCGGGGCCAATTCACCTCTGCAGGACAGAGGGGCGAGCCCGAGGGT 493
DB 264 GCCAGGTTCGAGTTGCGGATCGCATCTCGATGAGGCGCAGCTGGCTCGCTCGTAGGG 205
QY 494 GGAGATGATCTGCCAGGCGTCTCTCGGCGAGCCACCTATCACCAACAGCCTGATCGGGAA 553
DB 204 CTTGATCACCACCATGCGCGCTCGGGCAGCGTTGATGCTGGACAGCTCGTGTATCGGGGT 145
QY 554 GGATGGCGAGTCCACCTGCGAGCAGAGACCATGCCAGGCGAG 596
DB 144 GGAGGCGCGTAGTAGTGTGATCGGTTGACATGCCG 102

RESULT 21
US-08-997-362-179/c
Sequence 179, Application US/08997362
Patent No. 5985287
GENERAL INFORMATION:
APPLICANT: Tan, Paul
APPLICANT: Hiyama, Jun
APPLICANT: Visser, Elizabeth
APPLICANT: Skinner, Margot
APPLICANT: Scott, Linda
APPLICANT: Prestidge, Ross
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR
TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
NUMBER OF SEQUENCES: 194
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/997,362
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/873,970
FILING DATE: June 12, 1997
APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/705,347
FILING DATE: August 29, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1002c2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 179:
SEQUENCE CHARACTERISTICS:
LENGTH: 520 base pairs

SEQUENCE CHARACTERISTICS:
LENGTH: 520 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-997-362-179

Query Match
Best Local Similarity 3.8%; Score 37.4; DB 2; Length 520;
Matches 107; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 374 CACCTCAGTGCCTCCATGTGACAGTCCAGGCTACAGATGCACTGGGAGCTGTGTCCTCAA 433
DB 324 CTCCTCGGTGAGCTCGGGATCGACACCGGATGATGTCCTGCTGGTGGATTGAC 265
QY 434 GCCAGTGTCTGAGCTGCGGGGCCAATTCACCTCTGCAGGACAGAGGGGCGAGCCCGAGGGT 493
DB 264 GCCAGGTTCGAGTTGCGGATCGCATCTCGATGAGGCGCAGCTGGCTCGCTCGTAGGG 205
QY 494 GGAGATGATCTGCCAGGCGTCTCTCGGCGAGCCACCTATCACCAACAGCCTGATCGGGAA 553
DB 204 CTTGATCACCACCATGCGCGCTCGGGCAGCGTTGATGCTGGACAGCTCGTGTATCGGGGT 145
QY 554 GGATGGCGAGTCCACCTGCGAGCAGAGACCATGCCAGGCGAG 596
DB 144 GGAGGCGCGTAGTAGTGTGATCGGTTGACATGCCG 102

RESULT 22
US-09-095-855-179/c
Sequence 179, Application US/09095855
Patent No. 6160093
GENERAL INFORMATION:
APPLICANT: Tan, Paul
APPLICANT: Visser, Elizabeth
APPLICANT: Skinner, Margot
APPLICANT: Prestidge, Ross
TITLE OF INVENTION: Compounds and Methods for
Treatment and Diagnosis of Mycobacterial Infections
NUMBER OF SEQUENCES: 208
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/095,855
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/705,347
FILING DATE: 29-AUG-1996
APPLICATION NUMBER: 08/873,970
FILING DATE: 12-JUN-1997
APPLICATION NUMBER: 08/997,362
FILING DATE: 23-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1002c3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 179:
SEQUENCE CHARACTERISTICS:
LENGTH: 520 base pairs

```
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-095-855-179

Query Match      3.8%; Score 37.4; DB 3; Length 520;
Best Local Similarity 48.0%; Pred. No. 1.9;
Matches 107; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 374 CACCTCAGGTGCCATGTGGACAGTGCAGGCTACAGATGCACTGGAGCTGTGTCCTCAA 433
Db 324 CTCTCGGTGAGTGGGATGACACCCCGATGATGTTCCGCTGCTGCGATGAC 265
QY 434 GCCAGTGTCTGAGTGGCGGCCCACTTCACTCTGCAGGACAGAGGGGCGAGGCCCGAGGGT 493
Db 264 GCCAGTGGAGTGTGGATGCGATCTCTGATGAGGCGAGCTGCTGCCTCGTAGGG 205
QY 494 GGAGATGATCTGCAGGCGTCTCGGGCAGCCACCTATCACCACAGCCTGATCGGGAA 553
Db 204 CTTGATCACCACCATCGCGCTCGGGCAGCTTGTGATGTGACAGCTGCGTGTGATCGGGT 145
QY 554 GGATGGGCGAGTCCACCTCGCAGCAGACCATGCCACAGGCAG 596
Db 144 GGAGGCGCCGTAGTAGTGTGATGTCGGTGTGAACATGCCG 102

RESULT 23
US-09-324-542-179/c
; Sequence 179, Application US/09324542
; Patent No. 6328978
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L.J.
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; TITLE OF INVENTION: Of Immunologically-Mediated Skin Disorders
; FILE REFERENCE: 11000.100761
; CURRENT APPLICATION NUMBER: US/09/324,542
; CURRENT FILING DATE: 1999-06-02
; EARLIER APPLICATION NUMBER: US 08/997,080
; EARLIER FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 179
; LENGTH: 520
; TYPE: DNA
; ORGANISM: Mycobacterium vaccae
US-09-324-542-179

Query Match      3.8%; Score 37.4; DB 4; Length 520;
Best Local Similarity 48.0%; Pred. No. 1.9;
Matches 107; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 374 CACCTCAGGTGCCATGTGGACAGTGCAGGCTACAGATGCACTGGAGCTGTGTCCTCAA 433
Db 324 CTCTCGGTGAGTGGGATGACACCCCGATGATGTTCCGCTGCTGCGATGAC 265
QY 434 GCCAGTGTCTGAGTGGCGGCCCACTTCACTCTGCAGGACAGAGGGGCGAGGCCCGAGGGT 493
Db 264 GCCAGTGGAGTGTGGATGCGATCTCTGATGAGGCGAGCTGCTGCCTCGTAGGG 205
QY 494 GGAGATGATCTGCAGGCGTCTCGGGCAGCCACCTATCACCACAGCCTGATCGGGAA 553
Db 204 CTTGATCACCACCATCGCGCTCGGGCAGCTTGTGATGTGACAGCTGCGTGTGATCGGGT 145
QY 554 GGATGGGCGAGTCCACCTCGCAGCAGACCATGCCACAGGCAG 596
Db 144 GGAGGCGCCGTAGTAGTGTGATGTCGGTGTGAACATGCCG 102

RESULT 24
US-09-205-426-179/c
; Sequence 179, Application US/09205426
; Patent No. 6406704
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L. J.
; TITLE OF INVENTION: Compounds and Methods for Treatment and
; TITLE OF INVENTION: Diagnosis of Mycobacterial Infections
; FILE REFERENCE: 11000.100264
; CURRENT APPLICATION NUMBER: US/09/205,426
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: 09/095,855
; EARLIER FILING DATE: 1998-06-11
; EARLIER APPLICATION NUMBER: 08/997,362
; EARLIER FILING DATE: 1997-12-23
; EARLIER APPLICATION NUMBER: 08/873,970
; EARLIER FILING DATE: 1997-06-12
; EARLIER APPLICATION NUMBER: 08/705,347
; EARLIER FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 179
; LENGTH: 520
; TYPE: DNA
; ORGANISM: Mycobacterium vaccae
US-09-205-426-179

Query Match      3.8%; Score 37.4; DB 4; Length 520;
Best Local Similarity 48.0%; Pred. No. 1.9;
Matches 107; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 374 CACCTCAGGTGCCATGTGGACAGTGCAGGCTACAGATGCACTGGAGCTGTGTCCTCAA 433
Db 324 CTCTCGGTGAGTGGGATGACACCCCGATGATGTTCCGCTGCTGCGATGAC 265
QY 434 GCCAGTGTCTGAGTGGCGGCCCACTTCACTCTGCAGGACAGAGGGGCGAGGCCCGAGGGT 493
Db 264 GCCAGTGGAGTGTGGATGCGATCTCTGATGAGGCGAGCTGCTGCCTCGTAGGG 205
QY 494 GGAGATGATCTGCAGGCGTCTCGGGCAGCCACCTATCACCACAGCCTGATCGGGAA 553
Db 204 CTTGATCACCACCATCGCGCTCGGGCAGCTTGTGATGTGACAGCTGCGTGTGATCGGGT 145
QY 554 GGATGGGCGAGTCCACCTCGCAGCAGACCATGCCACAGGCAG 596
Db 144 GGAGGCGCCGTAGTAGTGTGATGTCGGTGTGAACATGCCG 102

RESULT 25
US-09-095-855-202/c
; Sequence 202, Application US/09095855
; Patent No. 6160093
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Compounds and Methods for
; TITLE OF INVENTION: Treatment and Diagnosis of Mycobacterial Infections
; NUMBER OF SEQUENCES: 208
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/095,855
; FILING DATE:
```

CLASSIFICATION:
PRIOR APPLICATION DATA: 08/705,347
APPLICATION NUMBER: 08/705,347
FILING DATE: 29-AUG-1996
APPLICATION NUMBER: 08/873,970
FILING DATE: 12-JUN-1997
APPLICATION NUMBER: 08/997,362
FILING DATE: 23-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1002c3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 202:
SEQUENCE CHARACTERISTICS:
LENGTH: 570 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-095-855-202

Query Match 3.8%; Score 37.4; DB 3; Length 570;
Best Local Similarity 48.0%; Pred. No. 2;
Matches 107; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 374 CACCTCAGGTGCGGAGTGGACAGTCCAGGCTACAGATGACCTGGGAGCTGTGTCCTCAA 433
DB 333 CTCCTCGGTGAGCTGCGGGATCGACACCGGATGATGTTGCGTCTGTCGATGAC 274
QY 434 GCGAGTGTCTGAGTGTGCGGCTCTCGGGCAGCCACCTATCACCACAGCTGATCGGGAA 553
DB 273 GCCAGGTGCGGATGCGGATCGATCTCATGAGGCGAGCTGGCTCGTATGAGG 214
QY 494 GGAGATGATCTGCCAGGCTCTCGGGCAGCCACCTATCACCACAGCTGATCGGGAA 553
DB 213 CTTGATCACCACCATGCGGCTCGGGCAGCTTGTATCGCGTTGACATGCGG 154
QY 554 GGATGGCAGGTCCACCTGCGAGCAGACCATGCCAGGCGAG 596
DB 153 GGAGGCGCGTAGTAGTGTGATGTTGATCGCGTTGACATGCGG 111

RESULT 26
US-09-205-426-202/c
Sequence 202, Application US/09205426
Patent No. 6406704
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Tan, Paul L. J.
TITLE OF INVENTION: Compounds and Methods for Treatment and
TITLE OF INVENTION: Diagnosis of Mycobacterial Infections
FILE REFERENCE: 11000.1002c4
CURRENT APPLICATION NUMBER: US/09/205,426
CURRENT FILING DATE: 1998-12-04
EARLIER APPLICATION NUMBER: 09/095,855
EARLIER FILING DATE: 1998-06-11
EARLIER APPLICATION NUMBER: 08/997,362
EARLIER FILING DATE: 1997-12-23
EARLIER APPLICATION NUMBER: 08/873,970
EARLIER FILING DATE: 1997-06-12
EARLIER APPLICATION NUMBER: 08/705,347
EARLIER FILING DATE: 1996-08-29
NUMBER OF SEQ ID NOS: 208
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 202
LENGTH: 570
TYPE: DNA
ORGANISM: Mycobacterium vaccae
US-09-205-426-202

Query Match 3.8%; Score 37.4; DB 4; Length 570;
Best Local Similarity 48.0%; Pred. No. 2;
Matches 107; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 374 CACCTCAGGTGCGGAGTGGACAGTCCAGGCTACAGATGACCTGGGAGCTGTGTCCTCAA 433
DB 333 CTCCTCGGTGAGCTGCGGGATCGACACCGGATGATGTTGCGTCTGTCGATGAC 274
QY 434 GCGAGTGTCTGAGTGTGCGGCTCTCGGGCAGCCACCTATCACCACAGCTGATCGGGAA 553
DB 273 GCCAGGTGCGGATGCGGATCGATCTCATGAGGCGAGCTGGCTCGTATGAGG 214
QY 494 GGAGATGATCTGCCAGGCTCTCGGGCAGCCACCTATCACCACAGCTGATCGGGAA 553
DB 213 CTTGATCACCACCATGCGGCTCGGGCAGCTTGTATCGCGTTGACATGCGG 154
QY 554 GGATGGCAGGTCCACCTGCGAGCAGACCATGCCAGGCGAG 596
DB 153 GGAGGCGCGTAGTAGTGTGATGTTGATCGCGTTGACATGCGG 111

RESULT 27
US-09-319-892-3
Sequence 3, Application US/09319892
Patent No. 6177616
GENERAL INFORMATION:
APPLICANT: BARTSCH, Klaus
APPLICANT: KRIETE, Guido
APPLICANT: BROER, Inge
APPLICANT: PUHLER, Alfred
TITLE OF INVENTION: NOVEL GENES CODING FOR AMINO ACID DEACETYLASES WITH
TITLE OF INVENTION: SPECIFICITY FOR N-ACETYL-L-PHOSPHINOTHRICIN, THEIR
TITLE OF INVENTION: ISOLATION AND THEIR USE
FILE REFERENCE: 514412-2005
CURRENT APPLICATION NUMBER: US/09/319,892
CURRENT FILING DATE: 1999-06-14
EARLIER APPLICATION NUMBER: PCT/EP97/06755
EARLIER FILING DATE: 1997-12-03
EARLIER APPLICATION NUMBER: 19652284.6
EARLIER FILING DATE: 1996-12-16
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 1273
TYPE: DNA
ORGANISM: Stenotrophomonas maltophilia
US-09-319-892-3

Query Match 3.7%; Score 36.6; DB 3; Length 1273;
Best Local Similarity 45.0%; Pred. No. 4.3;
Matches 138; Conservative 0; Mismatches 169; Indels 0; Gaps 0;

QY 431 CAAGCCAGTGTCTGAGCTGCGGSCCACTTCACTCTGAGGACAGAGGGGCGCCCGAG 490
DB 735 CGAGGCCACCAAGCTGTGCTGCCGACAGCGTGGAGTGGAGGCGCGTGGCAGCTTAC 794
QY 491 GGTGAGATGATCTGCCAGGCGTCTCTCGGGCAGCCACCTATCACCACAGCTGATCGG 550
DB 795 GCTGGAGGTGCTGGACCTGATCGAGCGGCGCATGAAGCCCTGGCGGAGAGCATCTGCC 854
QY 551 GAAGGATGGGAGGTCCACTGCGAGCAGACATGCGCAGGAGCTGCGCAACTTCTC 610
DB 855 GCGCATGACACGCGTGGAGTTCGAGTTCGTGCGCACTACCCGCCGCCACCATCACTC 914
QY 611 CTTCTGCGGAGCCAGACATCGGACTGTTCTGTTCTGTTGCGCAGGCTGCAAAACACCACTCT 670
DB 915 CGCCCGGAGGCGGATTCGACGCGCGCTCATGCGCGGAGTCTGTTGGGCGAGGCGCACT 974
QY 671 CCAGCAGCGCCCTCAGAGTGTGTCGCCCGGAGTGTGACAGAGATGAGGAGCTGGCA 730
DB 975 GCTGCCCCAGGAGCGCTCCATGGGCGCGGAGGACTTGGCTTTCATGCTGCTGAAAAGCC 1034

QY 731 GGGTCCC 737
Db 1035 CGGCGCC 1041

RESULT 28
US-09-727-238-3
; Sequence 3, Application US/09727238
; Patent No. 6555733
; GENERAL INFORMATION:
; APPLICANT: Bartsch, et al.
; TITLE OF INVENTION: No. 6555733el genes coding for amino acid deacetylases with speci
; FILE REFERENCE: acetyl-L-phosphinothricin, their isolation and their use.
; CURRENT APPLICATION NUMBER: US/09/727,238
; PRIOR FILING DATE: 2000-11-30
; PRIOR FILING DATE: 1999-06-14
; PRIOR FILING DATE: PCT/EP97/06755
; PRIOR FILING DATE: 1997-12-03
; PRIOR FILING DATE: 1996-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 1273
; TYPE: DNA
; ORGANISM: Comamonas acidovorans
; NAME/KEY: CDS
; LOCATION: (1)..(1206)
; OTHER INFORMATION: coding sequence of the deacetylase gene from comamonas acidovorans

US-09-727-238-3

Query Match 3.7%; Score 36.6; DB 4; Length 1273;
Best Local Similarity 45.0%; Pred. No. 4.3;
Matches 138; Conservative 0; Mismatches 169; Indels 0; Gaps 0;

QY 431 CAAGCCAGTCTTGAGCTGGCGGCCAATTCACCTCTGCAGACAGAGGGGCGAGCCCGCCAG 490
Db 735 CGAGGGCCACCAAGCTGTGCTGCCACAGCGTGGAGCTGCAGGGCAGGTGCGCACTTCCAC 794

QY 491 GGTGGAGATGATCTGCAGCGCTCTTCCGGCAGCCACCTATCACCAGCGCTGATCGG 550
Db 795 GCTGGAGGTGCTGGACCTGATGAGCGGGCGCATGAAAGGCCCTGGCGAGAGCATCTGCGC 854

QY 551 GAAGGATGGGCGAGGTCCACCTGCAGCAGAGACCATGCCAGCAGCGCTGCCAATCTTC 610
Db 855 GCGCATGACACGCGTGGAGTTCGAGTTCGTGCGCACTACCCGCCCACTCAACTC 914

QY 611 CTTCTCGCGAGCGCACATCGGACTGGTTGTGGTCCAGGCTGCACCAAGCCCAATGT 670
Db 915 CGCGCGAGGCGAGTTCGACGCGCGCTCATGGCGGAGGTGCTGGCGAGGCCAAGT 974

QY 671 CCAGCAGACGCGCTCACAGTGTGCCCCCGAGGTGTGACCAAGATGAGGACTGGCA 730
Db 975 GTGTCGCCAGAGCGCTCCATGGGCGCCGAGGACTTCGCTTCTGCTGAAAAGCC 1034

QY 731 GGGTCCC 737
Db 1035 CGGCGCC 1041

RESULT 29
US-09-252-991A-11102
; Sequence 11102, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11102
; LENGTH: 801
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11102

Query Match 3.7%; Score 36.4; DB 4; Length 801;
Best Local Similarity 50.0%; Pred. No. 4.1;
Matches 91; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 601 CCAACTTCTCTCTCTCTCCGAGCCAGACATCGGACTGTTCTGTGTCAGGCTGCAACA 660
Db 279 CCAAGGCTTTCCGCGAGCGCTCGACAAGCTGTGCGGATCAGCGTCAAGGAAGCAGAGG 338

QY 661 ACGCCATGTCCAGCAGCAGCGCTCTACAGTGTGCCCCCAGGTGTTGACCAAGATGG 720
Db 339 ACGGCGACATCTCTGCTCCGCTGCGCTGCTGCGCGCGCGCGGCAAGCAGATGATGG 398

QY 721 AGGACTGGCAGGCTCCCTGGAGAGCCCATCTTGCCTTTCGCGCTCTACAGGACACC 780
Db 399 TCAGCGCGCGCGCAGCGTGGCATCTCTCCCGCGAGCAACGCTGAACCTACAAGCCCT 458

QY 781 GC 782
Db 459 GC 460

RESULT 30
US-09-252-991A-11158
; Sequence 11158, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11158
; LENGTH: 1116
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11158

Query Match 3.7%; Score 36.4; DB 4; Length 1116;
Best Local Similarity 50.0%; Pred. No. 4.6;
Matches 91; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 601 CCAACTTCTCTCTCTCTCCGAGCCAGACATCGGACTGTTCTGTGTCAGGCTGCAACA 660
Db 683 CCAAGGCTTTCCGCGAGCGCTCGACAAGCTGTGCGGATCAGCGTCAAGGAAGCAGAGG 742

QY 661 ACGCCATGTCCAGCAGCAGCGCTCTACAGTGTGCCCCCAGGTGTTGACCAAGATGG 720
Db 743 ACGGCGACATCTCTGCTCCGCTGCGCTGCTGCGCGCGCGCGGCAAGCAGATGATGG 802

QY 721 AGGACTGGCAGGCTCCCTCGAGAGCCCATCTTGCCTTTCGCGCTCTACAGGACACC 780
Db 803 TCAGCGCGCGCGCAGCGTGGCATCTCTCCCGCGAGCAACGCTGAACCTACAAGCCCT 862

QY 781 GC 782
Db 459 GC 460

```
Db      863 GC 864

RESULT 31
US-09-252-991A-11094/C
; Sequence 11094, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11094
; LENGTH: 1584
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11094

Query Match      3.7%; Score 36.4; DB 4; Length 1584;
Best Local Similarity 50.0%; Pred. No. 5.2;
Matches 91; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 601 CCAACTTCTCTCTGCGAGCCAGACATCGACTGGTCTGTGTCGCCAGGCTGCAACA 660
Db      |||||
QY 839 CCAAGCCTTCCCGAGCGCCTCGACAGCTGTGCGGATCAGCGTCAAGGAAGCAGAGG 780
Db      |||||
QY 661 AGCCCAATGTCAGACAGCGCCCTCACAGTGGTGCGCCCGAGTGCGTGACCAAGAGATGG 720
Db      |||||
QY 779 ACGCGGACATCTGCTGCCGCTGCGCTGCTGCGCGCGCGGCGGACGAGATGTGG 720
Db      |||||
QY 721 AGGACTGGCAGGTCCTCTGGAGAGCCCATCTTGGCTTGGCGTCTTACAGGACGCC 780
Db      |||||
QY 719 TCGAGCGCGCGGCGACGCGTGCATCTCTCCCGGCGAGACGCTGAACACTACAGCCCT 660
Db      |||||
QY 781 GC 782
Db      ||
QY 659 GC 658

RESULT 32
US-09-211-930-12
; Sequence 12, Application US/09211930
; Patent No. 5962265
; GENERAL INFORMATION:
; APPLICANT: Tyrell E. No. 5962265ris
; APPLICANT: William Craig Moore
; APPLICANT: David Shay Silberstein
; TITLE OF INVENTION: HUMAN SIGNAL TRANSDUCTION
; FILE REFERENCE: PHM.70296
; CURRENT APPLICATION NUMBER: US/09/211,930
; CURRENT FILING DATE: 1998-12-15
; EARLIER APPLICATION NUMBER: GB 9726851.0
; EARLIER FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 2028
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-211-930-12

Query Match      3.7%; Score 36.4; DB 2; Length 2028;
Best Local Similarity 79.6%; Pred. No. 5.7;
Matches 43; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 893 AAAATGAGTGTGTTTAGCTCTTCCACAAAAA 992
Db      |||||
QY 1950 AAAACAAGTGTGTTTAAAGAGCTCCCGAAAAA 2003
Db      |||||

RESULT 33
US-09-340-993-12
; Sequence 12, Application US/09340993
; Patent No. 6034228
; GENERAL INFORMATION:
; APPLICANT: Tyrell E. No. 6034228ris
; APPLICANT: William Craig Moore
; APPLICANT: David Shay Silberstein
; TITLE OF INVENTION: HUMAN SIGNAL TRANSDUCTION SERINE/THREONINE KINASE
; FILE REFERENCE: PHM.70296.N1
; CURRENT APPLICATION NUMBER: US/09/340,993
; CURRENT FILING DATE: 1999-06-25
; EARLIER APPLICATION NUMBER: GB 9726851.0 & US 09/211,930
; EARLIER FILING DATE: 1997-12-19 & 1998-12-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 2028
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-340-993-12

Query Match      3.7%; Score 36.4; DB 3; Length 2028;
Best Local Similarity 79.6%; Pred. No. 5.7;
Matches 43; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 939 AAAATGAGTGTGTTTAGCTCTTCCACAAAAA 992
Db      |||||
QY 1950 AAAACAAGTGTGTTTAAAGAGCTCCCGAAAAA 2003
Db      |||||

RESULT 34
US-09-468-442-12
; Sequence 12, Application US/09468442
; Patent No. 6300098
; GENERAL INFORMATION:
; APPLICANT: Tyrell E. No. 6300098ris
; APPLICANT: William Craig Moore
; APPLICANT: David Shay Silberstein
; TITLE OF INVENTION: HUMAN SIGNAL TRANSDUCTION SERINE/THREONINE KINASE
; FILE REFERENCE: PHM.70296.N1
; CURRENT APPLICATION NUMBER: US/09/468,442
; CURRENT FILING DATE: 1999-12-21
; EARLIER APPLICATION NUMBER: US 09/340,993
; EARLIER FILING DATE: 1999-06-25
; EARLIER APPLICATION NUMBER: GB 9726851.0
; EARLIER FILING DATE: 1997-12-19
; EARLIER APPLICATION NUMBER: US 09/211,930
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 2028
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-468-442-12

Query Match      3.7%; Score 36.4; DB 4; Length 2028;
Best Local Similarity 79.6%; Pred. No. 5.7;
Matches 43; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 939 AAAATGAGTGTGTTTAGCTCTTCCACAAAAA 992
Db      |||||
QY 1950 AAAACAAGTGTGTTTAAAGAGCTCCCGAAAAA 2003
Db      |||||

RESULT 35
US-09-252-991A-5589/c

Query Match      3.7%; Score 36.4; DB 2; Length 2028;
Best Local Similarity 79.6%; Pred. No. 5.7;
Matches 43; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
```



```

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 2625
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-245-041-18

```

	Query Match	3.6%	Score 36;	DB 3;	Length 2625;
	Best Local Similarity	67.1%;	Pred. No. 7.9;		
	Matches 51;	Conservative 0;	Mismatches 25;	Indels 0;	Gaps 0;
QY	917	TGTTGCTATTGGAGTTCATGCAAAATGAGCTGCTGTTTTAGCTGCTCTGTGCCACAAAAA	976		
Db	2528	TGTTCTCTACTGCTCTTGAAATAAGGAGCAACATCCTTTGCTGCTGTGTAACAAAAA	2587		
QY	977	AAAAAAAAAAAAAAAAA	992		
Db	2588	AAAAAAAAAAAAAAAAA	2603		

```

RESULT 39
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

```

Query Match	3.6%	Score 36;	DB 3;	Length 4403765;
Best Local Similarity	47.9%;	Pred. No. 69;		
Matches 134;	Conservative 0;	Mismatches 145;	Indels 1;	Gaps 1;
QY 239	CTATTCCCTCTGTGGAACCAAGAACATCAAGTGGCCAGAGGTGGTGAAGACCCACGA	238		
Db	3773464	CGACGCCGCGTATTGGGCCAAGTACCGCGAGGATCGCGAAGTTCGGGCTGACCGAGGC		3773523
QY 299	GC CGGCCTCCTTCAACCTCAACCTTCACCTCAAGTCCAGTCCAGACCTGTCTCACTACTT	358		
Db	3773524	GATCGCGCGCTACAGCACCGCGCTGAAGATCACCCCGCGGTGTGGACGACGCCAC		3773583
QY 359	CTGCCGGGGTCTCCACCTCAGTGGCCATGTGGACAGTGCAGGCTACAGATGCAGT	418		
Db	3773584	GGGCTGAGCGGGCTGGCCCCCGCTCGCCGCCAGAGTGAATCCACGACGCGTTTCGGCG		3773643
QY 419	GGAGCTGTGTGTCACAGCCACAGTGTCTCAGCTGCGGGGCCAACTTCACTCTGCAGCACAGAGG	478		

```

Db      3773644  TGTGCGGTGCGCCGCTTTTCACTGTGCGGCGCAGAGGTTCCCGGAAGTC-GGCGGAGCGCGC 3773702
QY      479    GGCAGGGCCCCAGGGTGGAGATGATCTCCAGGCGTCTCTCG 518
Db      3773703  GCCGACCGCCAGGGTGGAGCGGCCATCTCCGTCGGCGACG 3773742

RESULT 40
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R. A. M.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

```

Query Match	3.6%;	Score 36;	DB 3;	Length 4411529;
Best Local Similarity	47.9%;	Pred. No. 69;		
Matches 134;	Conservative 0;	Mismatches 145;	Indels 1;	Gaps 1;
Qy	239	CTATTCCTCTGTGGAACTCAAGTGGCCAAAGGTGGTGAAGACCCACGA	298	
Db	3781282	CGACGCCGCTATTGGSCCAAGTACCGCAGGATGCGCGAAGTTCGGGCTGACCGAGGC	3781341	
Qy	299	GCGGGCTCCTTCAACCTCAACGTCACTCAAGTCCAGTCCAGACTGCTCACTACTT	358	
Db	3781342	GATCGCCGCTTACAGACACCCGGCTGAAGATCAACCCGACCTCGGTGTGGACGACGCCAC	3781401	
Qy	359	CTCGCGGGCTCTCCACTTCAGTGCCTATGTGGACAGTCCAGGCTACAGATGCAC	418	
Db	3781402	GGGCTGAGCGGGTGGCCCCCGCTGCGCGCAGATGAATCCACGACGCGTTTGGGGG	3781461	
Qy	419	GGAGCTGTGTGTCAGGCCAGTGTCTAGCTGCGGGCCAACTTCACTCTGCAGACAGG	478	
Db	3781462	TGTCGGTCCCGTTTCACTGTGCGGCGCAGAGTTTACCGBAAGTC-GCGCAGGCGCGC	3781520	
Qy	479	GGCAGGCCCCAGGGTGGAGATGATCTGCCAGGCGTCTCTCG	518	
Db	3781521	GCCGACCGCCAGGGTGAGCGGCCATCTCGTTCGGCGACG	3781560	

```

RESULT 41
US-08-258-261B-6/c
; Sequence 6, Application US/08258261B
; Patent No. 5639949
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaffney, Thomas Deane
; APPLICANT: Lam, Stephen Ting
; APPLICANT: Hammer, Philip E.
; APPLICANT: Uknes, Scott Joseph
; TITLE OF INVENTION: Genes for the synthesis of
; TITLE OF INVENTION: antipathogenic substances
; NUMBER OF SEQUENCES: 22

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/258,261B
FILING DATE: 08-JUN-1994
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/457,205
FILING DATE: 01-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-258-261B-6

Query Match 3.6%; Score 35.8; DB 1; Length 28958;
Best Local Similarity 52.3%; Pred. No. 21;
Matches 79; Conservative 0; Mismatches 72; Indels 0;

Qy	56	CCGAGGCATGGGCTCCCTGGCTGTTCTGCTTGGCGCTGCTGCTGCCAGAGCTTCTC	115
Db	4498	CAGGGCCATTGGCGCGGTGAGGCTTTGGCTCCGGCGCTCTTGTTGACGGCGGAGCTC	4439
Qy	116	CAAGGCACGGGAGGAGAAATTACCCCTGTGGTCTCCATTGCCCTACAAAGTCTCTGGAAGT	175
Db	4438	GAAGGACGGCGAGGACGGGATACCGTTTGGACGGCATCGAGAGCGCTCGAGCAGGA	4379
Qy	176	TTTCCCCAAGGCGCTGGGTGCTCAFAAC	206
Db	4378	GCATCCCGGCGCCTCGGCCACCGCGGAAC	4348

```

RESULT 42
US-08-456-837-6/c
/ Sequence 6, Application US/08456837
/ Patent No. 5643774
/ GENERAL INFORMATION:
/ APPLICANT: Schupp, Thomas
/ APPLICANT: Ligott, James M.
/ APPLICANT: Beck, James Joseph
/ APPLICANT: Hill, Dwight Steven
/ APPLICANT: Ryalls, John Andrew
/ APPLICANT: Gaffney, Thomas Deane
/ APPLICANT: Lam, Stephen Ting
/ APPLICANT: Hammer, Phillip E.
/ APPLICANT: Uknes, Scott Joseph
/ TITLE OF INVENTION: Genes for the synthesis of
/ TITLE OF INVENTION: antipathogenic substances
/ NUMBER OF SEQUENCES: 22
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Ciba-Geigy Corporation
/

```

STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/456,837
FILING DATE: 01-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-Jun-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-456-837-6

Query Match	3.6%	Score 35.8	DB 1	Length 28958	
Best Local Similarity	52.3%	Pred. No. 21			
Matches	79	Conservative 0	Mismatches 72	Indels 0	Gaps 0
QY	56	CCGAGGCATGGGGCTCCCTGGCGTCTCTGCTGGCGCGTCTGGCTGCCAGCAGCTTC	115		
Db	4498	CAGGGCCATTGGCGCGGTGAGGCCCTGGCTCCGGCGCGTCTGGTTCACGGCGGAGCCTC	4439		
QY	116	CAAGGCACGGGAGGAAGAATTACCCCTGTGGTCTCCATTGCCTACAAAGTCTCTGGAAGT	175		
Db	4438	GAAGGACGGCGAGACGGGATGACCGCTTTTGGACGGCATCGAGAGCGCGTCGAGCAGGA	4379		
QY	176	TTTTCCTCAAAGCGCGCTGGTGTCTATAACC	206		
Db	4378	GCATCCCGCGCGCTCGGCCACGCCGAACC	4348		

```

RESULT 43
US-08-457-342-6/c
/ Sequence 6, Application US/08457342
/ Patent No. 5662898
/ GENERAL INFORMATION:
/ APPLICANT: Schupp, Thomas M.
/ APPLICANT: Ligon, James M.
/ APPLICANT: Beck, James Joseph
/ APPLICANT: Hill, Dwight Steven
/ APPLICANT: Ryals, John Andrew
/ APPLICANT: Gaffney, Thomas Deane
/ APPLICANT: Lam, Stephen Ring
/ APPLICANT: Hammer, Phillip E.
/ APPLICANT: Uknes, Scott Joseph
/ TITLE OF INVENTION: Genes for the synthesis of
/ antipathogenic substances
/ NUMBER OF SEQUENCES: 22
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Ciba-Geigy Corporation

```


STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,342
FILING DATE: 01-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-Jun-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-457-342-6

	Query Match	3.6%;	Score 35.8;	DB 1;	Length 28958;
	Best Local Similarity	52.3%;	Pred. No. 21;		
	Matches	79;	Conservative	0;	Mismatches 72; Indels 0; Gaps 0;
QY	56	CCGAGGCATCGGCTCCCTCGGCTGTCTCTGTGGCGGTCTGGCTGCCACGACTTCTC			115
Db	4498	CAGGGCCATTGGGCGCGGTGAGGCTTTGGCTTCGGCGGCTCTGGTTCACGCGCGAGCCTC			4439
QY	116	CRAAGGCACGGGAGGAATAATCCCTGTGGTTCCTATTGCCTACAAAGTCCCTGGGAAGT			175
Db	4438	GAAGGACGGCGAGGACGGGATGACCGTTTGGACGGCATCGAGAGCGCTTCGAGCAGGA			4379
QY	176	TTTCCCGAAAGCGCGTGGGTGCTCTATAAC			206
Db	4378	GCATCCGGCGGCGCTCGGCCACGCCGAAC			4348

```

RESULT 44
US-08-457-646A-6/c
; Sequence 6, Application US/08457646A
; Patent No. 5679560
;
; GENERAL INFORMATION:
;
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaffney, Thomas Deane
; APPLICANT: Lam, Stephen Ting
; APPLICANT: Hammer, Phillip E.
; APPLICANT: Uknes, Scott Joseph
;
; TITLE OF INVENTION: Genes for the synthesis of
; TITLE OF INVENTION: antipathogenic substances
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
;

```

STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,646A
FILING DATE: 01-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-Jun-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-457-646A-6

	Query Match	3.6%	Score 35.8;	DB 1;	Length 28958;
	Best Local Similarity	52.3%;	Pred. No. 21;		
	Matches	79;	Conservative 0;	Mismatches 72;	Indels 0; Gaps 0;
Qy	56	CCGAGGCATCGGCTCCCTGGCGTGTCTGCTGGCGCGTCTGGCTGCCACGACGTTCTC	115		
Db	4498	CAGGCCATTGTTGGGCGCGTGGGCGCTTGGCTCGCGCGCTCTGGTTGACGCGGAGCGCTC	4439		
Qy	116	CAAGGCACGGGAGGAAGAAATATACCCCTGTGGTCTCCATTGCCTACAAAGTCCTCGAAGT	175		
Db	4438	GAAGACGGCGGAGGACGGATGACCGTTTGGACGGCATCGGAGCGGCTCGAGACGGA	4379		
Qy	176	TTTCCCAAAGGCGGCTGGTGTCTATAACC	206		
Db	4378	GCATCCCGGCGCGCTTCGGCCAGCGCCGAACC	4348		

```

RESULT 45
US-08-458-076A-6/c
? Sequence 6, Application US/08458076A
? Patent No. 5638425
? GENERAL INFORMATION:
? APPLICANT: Schupp, Thomas
? APPLICANT: Lagon, James M.
? APPLICANT: Beck, James Joseph
? APPLICANT: Hill, Dwight Steven
? APPLICANT: Ryals, John Andrew
? APPLICANT: Gaffney, Thomas Deane
? APPLICANT: Lam, Stephen Tiro
? APPLICANT: Hammer, Phillip E.
? APPLICANT: Uknes, Scott Joseph
? TITLE OF INVENTION: Genes for the synthesis of
? antipathogenic substances
? NUMBER OF SEQUENCES: 22
? CORRESPONDENCE ADDRESS:
? ADDRESSES: Ciba-Geigy Corporation

```



```
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,335A
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 800
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/457,205
; FILING DATE: 01-JUN-1995
; APPLICATION NUMBER: 08/258,261
; FILING DATE: 08-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-457-335A-6

Query Match          3.6%; Score 35.8; DB 1; Length 28958;
Best Local Similarity 52.3%; Pred. No. 21;
Matches 79; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 56 CCGAGGCGATGGGCTCCCTGGCGGTCTCTGCTGGCGGTCTGGCTGCCAGCAGCTTCTC 115
Db 4498 CAGGCGCATTTGGCGCGGTGAGGCTTGGCTCCGGCGGTCTGGTTGACGCGGAGCCTC 4439
QY 116 CAAGGCGCGGAGGAGAAATTACCCCTGGTCTCCATTCGCTTACAAAGTCCTTGAAGT 175
Db 4438 GAAGGCGCGGAGGAGCGGATGACCGTTTGGACGGCATCGAGAGCGGCTCGAGCAGGA 4379
QY 176 TTTCCCAAGGCGCGTGGGTGCTCATTAACC 206
Db 4378 GCATCCCGCGCCCTCGGCCCGAGCCGGAACC 4348

RESULT 48
US-08-729-214-6/c
; Sequence 6, Application US/08/29214
; Patent No. 5817502
; GENERAL INFORMATION:
; APPLICANT: Ligon, James M.
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Hammer, Phillip E.
; APPLICANT: van Pee, Karl-Heinz
; APPLICANT: Kirner, Sabine
; TITLE OF INVENTION: Genes for the synthesis of
; TITLE OF INVENTION: antipathogenic substances
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 520 White Plains Road
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA
```

```
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/729,214
; FILING DATE: TBA
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC 1506/CIP5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-729-214-6

Query Match          3.6%; Score 35.8; DB 1; Length 28958;
Best Local Similarity 52.3%; Pred. No. 21;
Matches 79; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 56 CCGAGGCGATGGGCTCCCTGGCGGTCTCTGCTGGCGGTCTGGCTGCCAGCAGCTTCTC 115
Db 4498 CAGGCGCATTTGGCGCGGTGAGGCTTGGCTCCGGCGGTCTGGTTGACGCGGAGCCTC 4439
QY 116 CAAGGCGCGGAGGAGAAATTACCCCTGGTCTCCATTCGCTTACAAAGTCCTTGAAGT 175
Db 4438 GAAGGCGCGGAGGAGCGGATGACCGTTTGGACGGCATCGAGAGCGGCTCGAGCAGGA 4379
QY 176 TTTCCCAAGGCGCGTGGGTGCTCATTAACC 206
Db 4378 GCATCCCGCGCCCTCGGCCCGAGCCGGAACC 4348

RESULT 49
US-09-028-934-6/c
; Sequence 6, Application US/09028934
; Patent No. 6117670
; GENERAL INFORMATION:
; APPLICANT: Ligon, James M.
; APPLICANT: Hill, Dwight S.
; APPLICANT: Lam, Steven T.
; APPLICANT: Hammer, Phillip E.
; APPLICANT: van Pee, Karl-Heinz
; APPLICANT: Kirner, Sabine
; APPLICANT: Young, Thomas R.
; TITLE OF INVENTION: Pyrolytic Biosynthesis Genes and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6117670artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```



```
; LOCATION: 40190..46318
; OTHER INFORMATION: /product= "Module 5 of SorB"
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 46851..47891
; OTHER INFORMATION: /product= "SorM"
; OTHER INFORMATION: /note= "The protein encoded by the sorM gene is highly
; OTHER INFORMATION: homologous to the methyltransferase from Streptomyces
; OTHER INFORMATION: hygroscopicus that is involved in the synthesis of the
; OTHER INFORMATION: polyketide rapamycin."
; US-08-764-233A-1
```

```
Query Match 3.6%; Score 35.8; DB 1; Length 49377;
Best Local Similarity 52.3%; Pred. No. 26;
Matches 79; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 56 CCGAGGCATGGGGCTCCCTGGGCTGTTCTGCTTGGCCGCTGGCTGGCCAGCAGCTTCTC 115
Db 20801 CAGGGCCATTGGGCGCGGTGAGGCCCTTGGCTCCGCGCTCTGTTGACGGCGGAGCCTC 20742

QY 116 CAAGCACGGGAGGAGAAATTACCCCTGTGTCTCCATTGCCTACAAAGTCCCTGGAAGT 175
Db 20741 GAAGGACGGCGAGGACGGGATACCGTTTGGACGGCATCGGAGAGGCGCTCGAGCAGGA 20682

QY 176 TTTCCCCAAAGGCCGCTGGGTGCTCATAACC 206
Db 20681 GCATCCCGCGCCCTCGGCCCGAGCCCGAACC 20651
```

Search completed: February 9, 2004, 15:01:36
Job time : 177 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2004, 12:59:37 ; Search time 328 Seconds
(without alignments)

8164.154 Million cell updates/sec

Title: US-09-990-726-222

Perfect score: 992

Sequence: 1 ggcagagccaggaactagg.....aaaaaaaaaaaaaaaaaaaaa 992

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 252756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : N_Geneseq_13Jun03.*

```

1: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	992	100.0	992	21	AAC58373 Human PRO809 nucle
2	992	100.0	992	22	AAF44176 Human PRO809 (UNQ4
3	992	100.0	992	24	ABK33610 cDNA encoding huma
4	992	100.0	992	25	ABX80267 Novel human secret
5	992	100.0	992	25	ABX80771 Human secreted/tra
6	992	100.0	992	25	ABX81154 Novel human secret
7	992	100.0	992	25	ABX90244 Human secreted/tra
8	992	100.0	992	25	ABX77855 Human PRO polynucl

9	992	100.0	992	25	ABX79451 Human secreted/tr
10	992	100.0	992	25	ABX64090 cDNA encoding huma
11	992	100.0	992	25	ABX17054 Human PRO polynucl
12	991	99.9	991	21	Membrane-bound pro
13	843	85.0	1047	22	Human secreted pro
14	790.2	79.7	935	22	Human EST-derived
15	773.6	78.0	1047	21	Human secreted pro
16	773.6	78.0	1047	25	ABT16794 Human secreted pro
17	773.6	78.0	1047	25	ABZ66993 Human secreted pro
18	501.6	50.6	502	25	ABZ36822 Human GENSET codin
19	501.6	50.6	724	22	Human cDNA encodin
20	292.8	29.5	3220	22	Human nervous syst
21	149.6	15.1	150	19	Human biallelic po
22	149.6	15.1	150	19	Human biallelic po
23	47.2	4.8	1272	24	Human NOV93a cDNA.
24	47	4.7	53226	25	Human G-protein co
25	47	4.7	87350	18	Human WRN genomic
26	44.4	4.5	636	23	DNA encoding novel
27	44	4.4	249999	25	Human tramdorin ge
28	43.4	4.4	560	24	Human breast cance
29	43	4.3	509	24	Human breast cance
30	42.8	4.3	143899	24	Genomic sequence e
31	41.8	4.2	752	21	Human breast and o
32	41.8	4.2	1417	21	Human ORFX ORF2848
33	41.8	4.2	2059	21	Human cancer assoc
34	41.8	4.2	2672	25	Human cDNA #699 di
35	41.8	4.2	3471	22	Human cDNA sequenc
36	41.8	4.2	3541	24	Human poly(ADP-rib
37	41	4.1	38736	24	Human membrane spa
38	39.6	4.0	44100	21	Human retroviral s
39	39.4	4.0	276	24	Human ORFX polynucl
40	39.4	4.0	3285	20	Clone ml03.4 enco
41	39.4	4.0	3285	22	Human cDNA encodin
42	39.4	4.0	3285	22	Human polynucl
43	39.2	4.0	394	22	Novel human polynu
44	38.8	3.9	1639	21	Human colon cancer
45	38.8	3.9	2025	24	cDNA encoding huma
46	38.6	3.9	56737	24	Human hypoxanthine
47	38.4	3.9	185695	25	Human prostate exp
48	38	3.8	294	23	Bifidobacterium lo
49	38	3.8	349980	24	Human prostate exp
50	37.8	3.8	353	23	Myobacterium tube
51	37.8	3.8	4403765	22	Myobacterium tube
52	37.8	3.8	4411529	22	Myobacterium tube
53	37.6	3.8	375	22	Human polynucl
54	37.6	3.8	861	24	Human polynucl
55	37.6	3.8	1048	21	cDNA encoding a ma
56	37.6	3.8	1396	16	Human 50 kDa dystr
57	37.6	3.8	1396	18	Human (50 kDa) dys
58	37.6	3.8	1602	24	Human polynucl
59	37.6	3.8	1626	25	High-affinity phos
60	37.6	3.8	1626	25	High-affinity Osph
61	37.6	3.8	3083	25	Human gene sequenc
62	37.6	3.8	3537	25	Human secreted pro
63	37.6	3.8	80959	25	Human secreted pro
64	37.4	3.8	520	20	Nucleotide sequenc
65	37.4	3.8	520	24	M vaccae GV-41 cod
66	37.4	3.8	570	20	Nucleotide sequenc
67	37.4	3.8	3675	21	Human breast and o
68	37.2	3.8	296	22	Human breast cance
69	37.2	3.8	450	19	Cysteine protease
70	37	3.7	697	23	Drosophila melanog
71	36.8	3.7	512	24	Human ovarian can
72	36.8	3.7	545	24	Human colon cancer
73	36.8	3.7	784	23	ABQ58601 Human
74	36.8	3.7	2693	23	ABL08343 Drosophila melanog
75	36.8	3.7	2784	23	ABL11520 Drosophila melanog
76	36.6	3.7	1273	19	ABL08342 Drosophila melanog
77	36.6	3.7	1571	20	ABL45110 Comamonas sp. N-AC
78	36.4	3.7	381	20	ABL90445 Human polynucl
79	36.4	3.7	381	20	ABL80803 Human secreted/tra
80	36.4	3.7	745	19	Novel human secret
81	36.4	3.7	745	19	Human secreted/tra
					Clone dn809.5 isol
					Human polynucl

C 82	36.4	3.7	760	20	AAAG1427	DNA encoding a hum
C 83	36.4	3.7	760	25	ACG50346	Human secreted pro
C 84	36.4	3.7	760	25	ABZ71191	Human secreted pro
C 85	36.4	3.7	809	21	AACT79824	Human secreted pro
C 86	36.4	3.7	2028	20	AAAG1861	cDNA encoding a mu
C 87	36.4	3.7	2028	21	AAAG08010	Mouse polynucleoti
C 88	36.4	3.7	2044	20	AAV84372	Human epidermoid c
C 89	36.2	3.6	295	22	AAZ29052	cDNA encoding for
C 90	36.2	3.6	295	24	ABG68192	cDNA encoding huma
C 91	36.2	3.6	304	22	AAZ29128	cDNA encoding for
C 92	36.2	3.6	304	24	ABG68268	cDNA encoding huma
C 93	36.2	3.6	359	22	AAZ23694	Human breast cance
C 94	36.2	3.6	590	24	ABL59497	EST related to hum
C 95	36.2	3.6	1641	20	AAZ10823	Choline oxidase (C
C 96	36.2	3.6	2821	22	AAI66075	Triticum aestivum
C 97	36.2	3.6	2920	20	ABL35012	Rat cDNA isolated
C 98	36.2	3.6	541	20	AAAG7415	Repatocellular car
C 99	36.2	3.6	839	22	AAAG6484	Maize Mac20 coding
C 100	36.2	3.6	1127	22	AAAD07899	Human secreted pro
C 101	36.2	3.6	1128	22	AAAD07835	Human secreted pro
C 102	36.2	3.6	1805	21	AAZ58260	Soybean histone de
C 103	36.2	3.6	1975	24	ABL90817	Human polynucleoti
C 104	36.2	3.6	2101	21	AAZ93390	Human secreted pro
C 105	36.2	3.6	2101	22	AAZ32713	Human secreted pro
C 106	36.2	3.6	2101	25	ABZ73609	Secreted protein-e
C 107	36.2	3.6	2101	25	ABZ67203	Human secreted pro
C 108	36.2	3.6	2625	21	AAZ31922	Human mahogany pro
C 109	36.2	3.6	2626	16	AAQ94582	Modified Tag DNA-p
C 110	36.2	3.6	10579	22	ABA20014	Human nervous syst
C 111	36.2	3.6	10579	21	AAAG70045	Human immune/haema
C 112	36.2	3.6	21185	21	AAAG63350	Streptomyces globi
C 113	36.2	3.6	28995	22	ABZ20015	Human nervous syst
C 114	36.2	3.6	28995	22	AAK70046	Human immune/haema
C 115	36.2	3.6	28995	22	AAK79967	Human immune/haema
C 116	36.2	3.6	28995	22	AAK85213	Human immune/haema
C 117	36.2	3.6	63164	21	AAAG63348	Streptomyces globi
C 118	36.2	3.6	4403765	22	AAI99683	Mycobacterium tube
C 119	36.2	3.6	4411529	22	AAI99682	Mycobacterium tube
C 120	35.8	3.6	879	24	ABN98766	Arabidopsis thalia
C 121	35.8	3.6	1132	21	AAZ76120	Human ORFX ORF1675
C 122	35.8	3.6	1725	24	ABA95144	Human lysyl-oxidas
C 123	35.8	3.6	2328	24	AAI67786	Human lysyl oxidas
C 124	35.8	3.6	28598	17	AAI06769	Sorangium cellulos
C 125	35.8	3.6	28598	18	AAI99956	Sorangium cellulos
C 126	35.8	3.6	28958	21	AAZ75299	DNA sequence of So
C 127	35.8	3.6	43377	19	AAV05287	The soraphen biosy
C 128	35.8	3.6	144460	21	AAZ93815	Olfactory receptor
C 129	35.8	3.6	1503841	24	ABT00010	Human neuregulin 1
C 130	35.8	3.6	1503841	24	ABT01503	Human neuregulin 1
C 131	35.8	3.6	1503900	22	AAK95240	Human neuregulin-1
C 132	35.8	3.6	1503900	22	AAK96733	Human neuregulin-1
C 133	35.6	3.6	366	21	AAZ92717	Human wild-type cy
C 134	35.6	3.6	387	21	AAZ92716	Human wild-type cy
C 135	35.6	3.6	500	21	AAZ94277	Arabidopsis thalia
C 136	35.6	3.6	577	24	ABL66282	Lung cancer relate
C 137	35.6	3.6	577	24	ABL67527	Thyroid cancer rel
C 138	35.6	3.6	577	24	ABL67787	Oesophagus cancer
C 139	35.6	3.6	581	21	AAZ35658	Human cystatin E
C 140	35.6	3.6	588	21	AAZ60085	Human cystatin E
C 141	35.6	3.6	588	22	AAZ20097	Human cystatin E
C 142	35.6	3.6	588	24	AAZ43452	Human gene express
C 143	35.6	3.6	591	24	AAZ35213	Human cysteine pro
C 144	35.6	3.6	598	18	AAZ60608	Pseudomonas aerugi
C 145	35.6	3.6	1536	23	AAZ51510	Human ovarian anti
C 146	35.6	3.6	2046	24	ABQ54507	Human lysosomal ac
C 147	35.6	3.6	16759	24	ABL36297	Human prostate exp
C 148	35.4	3.6	294	23	ABV61758	Human prostate exp
C 149	35.4	3.6	356	23	ABV10788	Human prostate exp
C 150	35.4	3.6	361	23	ABV01619	Human prostate exp

ALIGNMENTS

RESULT 1

AAC58373 standard; cDNA; 992 BP.

XX AAC58373;

XX AAC58373;

DT 29-JAN-2001 (first entry)

XX Human PRO809 nucleotide sequence SEQ ID NO:22.

Human; tumour; diagnosis; neoplastic disease; neoplastic cell growth; proliferation; tumorigenesis; identification; cancer; cytostatic; neotropic; neuroprotective; antiinflammatory; immunosuppressive; immunostimulant; antiangiogenic; leukaemia; lymphoid malignancy; neuronal disorder; glial disorder; astrocytal disorder; angiogenic; hypothalamic disorder; glandular disorder; macrophagal disorder; epithelial disorder; stromal disorder; blastocoelec disorder; inflammatory disorder; immunologic disorder; ss.

XX Homo sapiens.

XX WO200053755-A2.

XX 14-SEP-2000.

XX 06-JAN-2000; 2000WO-US00376.

XX 08-MAR-1999; 99WO-US05028.

XX 02-JUN-1999; 99WO-US12252.

XX 23-JUN-1999; 99US-0141037.

XX 07-JUL-1999; 99US-0143048.

XX 26-JUL-1999; 99US-0145698.

XX 30-NOV-1999; 99WO-US28313.

XX 20-DEC-1999; 99WO-US30911.

XX 05-JAN-2000; 2000WO-US00219.

(GETH) GENENTECH INC.

XX Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hillan KJ, Roy MA; Watanabe CK, Wood WI;

XX WPI; 2000-572270/53.

XX P-PSDB; AAB24063.

XX Thirty PRO polynucleotides encoding PRO polypeptides, useful in the treatment, diagnosis and prevention of cancer -

XX Claim 50; Fig 13; 286pp; English.

The present invention describes an isolated antibody that binds to one of the human PRO proteins designated PRO212, PRO290, PRO341, PRO535, PRO619, PRO717, PRO809, PRO830, PRO848, PRO943, PRO1005, PRO1009, PRO1025, PRO1030, PRO1097, PRO1107, PRO1111, PRO1153, PRO1182, PRO1184, PRO1187, PRO1281, PRO23, PRO39, PRO834, PRO1317, PRO1710, PRO2094, PRO2145 OR PRO2198. PRO antagonists can be used to inhibit tumour cell growth. The PRO polypeptides and nucleotides are useful in the treatment, diagnosis and prevention of cancer. The antibodies and other anti-tumour compounds may be used to treat various conditions, including those characterised by overexpression and/or activation of the amplified PRO genes. Exemplary conditions or disorders to be treated with such antibodies and other compounds include benign or malignant tumours (e.g., renal, liver, kidney, bladder, breast, gastric, thyroid, ovarian, colorectal, prostate, pancreatic, lung, vulva, thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and neck tumours), leukaemias and lymphoid malignancies, other disorders such as neuronal, glial, astrocytal, hypothalamic and other glandular, macrophagal, epithelial, stromal and blastocoelec disorders, and inflammatory, angiogenic and immunologic disorders. AAC58242 to AAC58366 represent PCR primers and hybridisation probes used in the isolation of the human PRO sequences. AAC58367 to AAC58396 and AAB24057 to AAB24089 represent human PRO polynucleotide and protein sequences given in the exemplification of the present invention.

CC and DNA. They may also be used to produce transgenic animals which are
CC used to develop and screen therapeutically useful reagents. The PRO
CC nucleotide and protein sequence can be used for tissue typing and in
CC treating cancer. Anti-PRO antibodies can be used in diagnostic assays.
CC AAF44270 to AAF44470 represent PCR primers and hybridisation probes used
CC in the isolation of human PRO sequences. AAF44087 to AAF44269 and
CC AAB65154 to AAB65300 represent human PRO polynucleotide and protein
CC sequences given in the exemplification of the present invention.
XX
SQ Sequence 992 BP; 235 A; 301 C; 282 G; 174 T; 0 other;

Query Match 100.0%; Score 992; DB 22; Length 992;
Best Local Similarity 100.0%; Pred. No. 2.7e-236; Indels 0; Gaps 0;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACAGCCAGGAACCTAGGAGTTCTCACTGCCGAGCAGAGGCCCTACACCCACCGAG 60
DB 1 GGCACAGCCAGGAACCTAGGAGTTCTCACTGCCGAGCAGAGGCCCTACACCCACCGAG 60

QY 61 GCATGGGGTCCCTGGCTGTTCTGCTGGCGGTGCTGCTGCCAGCAGCTTCTCCAGG 120
DB 61 GCATGGGGTCCCTGGCTGTTCTGCTGGCGGTGCTGCTGCCAGCAGCTTCTCCAGG 120

QY 121 CACGGAGGAGAAATTACCCCTGTGTCCTCAATGCCCTACAAAGTCTCGAAGTTTTC 180
DB 121 CACGGAGGAGAAATTACCCCTGTGTCCTCAATGCCCTACAAAGTCTCGAAGTTTTC 180

QY 181 CCAAGGGCGGTGGTGTCTATACCTGTGTGACCCAGCCAGCCAGCCCATCACT 240
DB 181 CCAAGGGCGGTGGTGTCTATACCTGTGTGACCCAGCCAGCCAGCCCATCACT 240

QY 241 ATTCCCTCTGTGGAACCAAGAACATCAAGGTGGCCAGAGTGTGTGAAGCCCAAGC 300
DB 241 ATTCCCTCTGTGGAACCAAGAACATCAAGGTGGCCAGAGTGTGTGAAGCCCAAGC 300

QY 301 CGGCTCTCTCAACCTCAACCTCAACCTCAACCTCAACCTCAACCTCAACCTCACT 360
DB 301 CGGCTCTCTCAACCTCAACCTCAACCTCAACCTCAACCTCAACCTCAACCTCACT 360

QY 361 GCGGGGCTCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTGG 420
DB 361 GCGGGGCTCTCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTGG 420

QY 421 AGCTGTGTCTCAAGCCAGTGTCTGAGTGGCGGCCAATCTCACTCTGCAGCAGAGGG 480
DB 421 AGCTGTGTCTCAAGCCAGTGTCTGAGTGGCGGCCAATCTCACTCTGCAGCAGAGGG 480

QY 481 CAGGCCCCAGGGTGAGATGATCTGCCAGGGCTCTCGGGCAGGCCCACTATCAACCA 540
DB 481 CAGGCCCCAGGGTGAGATGATCTGCCAGGGCTCTCGGGCAGGCCCACTATCAACCA 540

QY 541 GCCTGATCGGAGAGATGGGAGGTCTCACTCTGCAGCAGAGACCATGCAAGGAGCTGG 600
DB 541 GCCTGATCGGAGAGATGGGAGGTCTCACTCTGCAGCAGAGACCATGCAAGGAGCTGG 600

QY 601 CCAACTCTCTCTCTCTCGGAGCAGAGATCGAGTCTGCTGCTGCTGCTGCTGCTGCTG 660
DB 601 CCAACTCTCTCTCTCTCGGAGCAGAGATCGAGTCTGCTGCTGCTGCTGCTGCTGCTG 660

QY 661 ACGCCAAATGTCCAGCAGAGCCCTCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 720
DB 661 ACGCCAAATGTCCAGCAGAGCCCTCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 720

QY 721 AGGACTGCGAGGTCTCTGAGAGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
DB 721 AGGACTGCGAGGTCTCTGAGAGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780

QY 781 GCGGTCTGAGTGAAGAGAGTTTGGGGGTTTCAGGATAGGGAATGGGGAGTCTCAGGAC 840
DB 781 GCGGTCTGAGTGAAGAGAGTTTGGGGGTTTCAGGATAGGGAATGGGGAGTCTCAGGAC 840

QY 841 GCAAGCAGCAGCCATGTAGTAATGAACCTGCAGAGAGCCAGCAGCCAGGAGGACTGCA 900
DB 841 GCAAGCAGCAGCCATGTAGTAATGAACCTGCAGAGAGCCAGCAGCCAGGAGGACTGCA 900

Db 841 GCAAGCAGCAGCCATGTAGTAATGAACCTGCAGAGAGCCAGCAGGAGGACTGCA 900

QY 901 GGCATCAGCGCTGCACCTGTTCTGTTATTTGGAGTTTCATGCAAAATGAGTGTGTTTGTAGCTGC 960
DB 901 GGCATCAGCGCTGCACCTGTTCTGTTATTTGGAGTTTCATGCAAAATGAGTGTGTTTGTAGCTGC 960

QY 961 TCTTGGCCACAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 992
DB 961 TCTTGGCCACAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 992

RESULT 3
ABK3610
ID ABK3610 standard; cDNA; 992 BP.
XX
AC ABK3610;
XX
DT 08-MAY-2002 (first entry)
XX
DE cDNA encoding human PRO protein, Seq ID No 149.
XX
KW Human; secreted protein; PRO; tumour; lung cancer; colon cancer;
KW breast cancer; prostate tumour; rectal tumour; liver tumour;
KW pericyte cell proliferation; chondrocyte cell proliferation;
KW tumour necrosis factor-alpha; gene; ss.
XX
OS Homo sapiens.
XX
FN WO200208288-A2.
XX
PD 31-JAN-2002.
XX
PF 29-JUN-2001; 2001WO-US21066.
XX
PR 20-JUL-2000; 2000US-219556P.
PR 25-JUL-2000; 2000US-220585P.
PR 25-JUL-2000; 2000US-220605P.
PR 25-JUL-2000; 2000US-220607P.
PR 25-JUL-2000; 2000US-220624P.
PR 25-JUL-2000; 2000US-220638P.
PR 25-JUL-2000; 2000US-220664P.
PR 26-JUL-2000; 2000US-220893P.
PR 28-JUL-2000; 2000WO-US20710.
PR 23-AUG-2000; 2000WO-US23522.
PR 24-AUG-2000; 2000WO-US23328.
PR 15-SEP-2000; 2000US-00000P.
PR 10-NOV-2000; 2000WO-US30873.
PR 28-NOV-2000; 2000US-253646P.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000US-074259.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 10-MAY-2001; 2001US-0854280.
PR 25-MAY-2001; 2001WO-US17092.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PU;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WT;
XX
WI MPI; 2002-172001/22.
DR P-PSDB; AAU83666.
XX
PT One hundred and twenty two nucleic acids encoding PRO polypeptides,
PT useful for treating a PRO related disorder and for diagnosing tumours
PT such as lung cancer, colon cancer, breast tumour, prostate tumour, rectal
PT tumour or liver tumour -
XX
PS Claim 2; Figure 149; 359pp; English.
XX
CC The invention relates to one hundred and twenty two nucleic acids
CC encoding PRO polypeptides. The sequences of the 122 PRO polynucleotides

CC encode human secreted proteins. The PRO nucleic acids, polypeptides, agonists and antagonists are useful for treating a PRO related disorder.
 CC The PRO polypeptides are useful for diagnosing tumours, especially lung cancer, colon cancer, breast tumour, prostate tumour, rectal tumour or liver tumour. The PRO polypeptides are useful for stimulating the proliferation of, or gene expression, in pericyte cells, for stimulating the proliferation or differentiation of chondrocyte cells, for stimulating the release of tumour necrosis factor-alpha from human blood, for stimulating or inhibiting the proliferation of normal human dermal fibroblast cells. The PRO polypeptide may also be used as molecular weight markers and for tissue typing. The PRO nucleic acids have applications in molecular biology, including use as hybridisation probes, CC and in chromosome and gene mapping. ABK33536-ABK33657 represent human CC PRO protein coding sequences of the invention.
 XX
 SQ

Sequence 992 BP; 235 A; 301 C; 282 G; 174 T; 0 other;

Query Match 100.0%; Score 992; DB 24; Length 992;
 Best Local Similarity 100.0%; Pred. No. 2.7e-236;
 Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACGAGCCAGCACTAGGAGTTCTCACTCCGAGCAGAGCCCTACCCACCAG 60
 Db 1 GGCACGAGCCAGCACTAGGAGTTCTCACTCCGAGCAGAGCCCTACCCACCAG 60
 QY 61 GCATGGGGCTCCCTGGGCTGTCTGCTGGCGTGGCTGCCAGCAGCTTCTCAAG 120
 Db 61 GCATGGGGCTCCCTGGGCTGTCTGCTGGCGTGGCTGCCAGCAGCTTCTCAAG 120
 QY 121 CACGGAGGAGAAATTAACCCCTGTGCTCCATTCGCTCAAAAGTCTCTGGAAGTTTCC 180
 Db 121 CACGGAGGAGAAATTAACCCCTGTGCTCCATTCGCTCAAAAGTCTCTGGAAGTTTCC 180
 QY 181 CCAAGGCCCGTGGTCTCTATACCTGTGTGACACCCAGCAGCACCAGCCATCACT 240
 Db 181 CCAAGGCCCGTGGTCTCTATACCTGTGTGACACCCAGCAGCACCAGCCATCACT 240
 QY 241 ATTCCCTCTCTGGAACCAAGAAATCAATCAAGTGGCAGAGAGTGGTGAACACCAAGC 300
 Db 241 ATTCCCTCTCTGGAACCAAGAAATCAATCAAGTGGCAGAGAGTGGTGAACACCAAGC 300
 QY 301 CGGCTCTCTTCAACTCAAGTCACTCACTCAAGTCCAGTCCAGACCTGCTCACCTACTTCT 360
 Db 301 CGGCTCTCTTCAACTCAAGTCACTCACTCAAGTCCAGTCCAGACCTGCTCACCTACTTCT 360
 QY 361 GCGGGGCTCCCTCACTCAGTGGCCATGTGACAGTGCAGGCTACAGATGCACCTGGG 420
 Db 361 GCGGGGCTCCCTCACTCAGTGGCCATGTGACAGTGCAGGCTACAGATGCACCTGGG 420
 QY 421 AGCTGTGGTCCAAAGCCAGTGTCTGAGCTGGGGCCAACTTCACTCTGAGCAGAGGGGG 480
 Db 421 AGCTGTGGTCCAAAGCCAGTGTCTGAGCTGGGGCCAACTTCACTCTGAGCAGAGGGGG 480
 QY 481 CAGGCCCGAGGTGGAGATGATCTGCAGCGCTCTCGGGCAGCCGCTATCAACCA 540
 Db 481 CAGGCCCGAGGTGGAGATGATCTGCAGCGCTCTCGGGCAGCCGCTATCAACCA 540
 QY 541 GCCTGATCGGGAAGGATGGCAGTCCACCTGAGCAGAGACCATGCACAGCAGCCTG 600
 Db 541 GCCTGATCGGGAAGGATGGCAGTCCACCTGAGCAGAGACCATGCACAGCAGCCTG 600
 QY 601 CCAACTTCTCTTCTCCGAGCAGACATCGGACTGGTTCTGGTCCAGGCTGCAACA 660
 Db 601 CCAACTTCTCTTCTCCGAGCAGACATCGGACTGGTTCTGGTCCAGGCTGCAACA 660
 QY 661 ACGCCAATGTCAGCAGCGCCCTCACAGTGTGCCCCAGTGTGTGACCAAGATGG 720
 Db 661 ACGCCAATGTCAGCAGCGCCCTCACAGTGTGCCCCAGTGTGTGACCAAGATGG 720
 QY 721 AGGACTGGCAGGCTCCCTGGAGAGCCCATCTTGTGCTTGCCTCTACAGAGCACCC 780
 Db 721 AGGACTGGCAGGCTCCCTGGAGAGCCCATCTTGTGCTTGCCTCTACAGAGCACCC 780

QY 781 GCCGTCTGAGTGAAGAGGAGTTTGGGGGTTTCAGGATAGGAATGGGAGTTCAGAGAC 840
 Db 781 GCCGTCTGAGTGAAGAGGAGTTTGGGGGTTTCAGGATAGGAATGGGAGTTCAGAGAC 840
 QY 841 GCAAGCAGCAGCAGCATGTAGAAATGAACCGTCCAGAGAGCCAAAGCAGCGCAGAGCTGCA 900
 Db 841 GCAAGCAGCAGCAGCATGTAGAAATGAACCGTCCAGAGAGCCAAAGCAGCGCAGAGCTGCA 900
 QY 901 GGCATCAGCGTGCACCTGTTCGTTATTTGGAGTTTCATGCAAAATGAGTGTGTTTAGCTGC 960
 Db 901 GGCATCAGCGTGCACCTGTTCGTTATTTGGAGTTTCATGCAAAATGAGTGTGTTTAGCTGC 960
 QY 961 TCTTGGCCACAAAAAATAAAAAAAAAAAAAAAAAAAAAA 992
 Db 961 TCTTGGCCACAAAAAATAAAAAAAAAAAAAAAAAAAAAA 992

RESULT 4

ABX80267
 ID ABX80267 standard; DNA; 992 BP.
 AC ABX80267;
 XX
 DT 28-APR-2003 (first entry)
 XX
 DE Novel human secreted or transmembrane protein PRO791 DNA.

XX Human; PRO; hypertrophy of neonatal heart; angiogenesis; wound healing;
 KW cardiac insufficiency disorder; cancer; tumour; immune response;
 KW adrenal cortical capillary endothelial growth; c-fos induction;
 KW vascular endothelial growth factor inhibition; VEGF inhibition;
 KW endothelial cell growth inhibitor; T-lymphocytes stimulation;
 KW retinal neurons cell survival; rod photoreceptor cell survival;
 KW retinal disorder; retinitis pigmentosa; kidney disease;
 KW mammalian kidney mesangial cell proliferation; Berger disease;
 KW dermatitis; herpetic keratitis; Crohn's disease; chondrocyte proliferation;
 KW chondrocyte redifferentiation; sports injury; arthritis; gene; ds.

OS Homo sapiens.

XX US2002132252-A1.

XX 19-SEP-2002.

XX 14-NOV-2001; 2001US-0990442.

XX 05-NOV-1997; 97WO-US200069.

PR 16-SEP-1998; 98WO-US19330.

PR 17-SEP-1998; 98WO-US19437.

PR 07-OCT-1998; 98WO-US21141.

PR 01-DEC-1998; 98WO-US25108.

PR 05-JAN-1999; 99WO-US00106.

PR 08-MAR-1999; 99WO-US05028.

PR 15-SEP-1999; 99WO-US21090.

PR 15-SEP-1999; 99WO-US21547.

PR 30-NOV-1999; 99WO-US28313.

PR 01-DEC-1999; 99WO-US28301.

PR 16-DEC-1999; 99WO-US28634.

PR 20-DEC-1999; 99WO-US30095.

PR 06-JAN-2000; 99WO-US30911.

PR 06-JAN-2000; 2000WO-US00219.

PR 06-JAN-2000; 2000WO-US00376.

PR 11-FEB-2000; 2000WO-US03565.

PR 18-FEB-2000; 2000WO-US04341.

PR 22-FEB-2000; 2000WO-US04414.

PR 24-FEB-2000; 2000WO-US04914.

PR 24-FEB-2000; 2000WO-US05004.

PR 02-MAR-2000; 2000WO-US05841.

PR 10-MAR-2000; 2000WO-US06319.

PR 15-MAR-2000; 2000WO-US06884.

PR 20-MAR-2000; 2000WO-US07377.

PR 30-MAR-2000; 2000WO-US08439.

PR 15-MAY-2000; 2000WO-US13358.
 PR 17-MAY-2000; 2000WO-US13705.
 PR 22-MAY-2000; 2000WO-US14042.
 PR 30-MAY-2000; 2000WO-US14941.
 PR 02-JUN-2000; 2000WO-US15264.
 PR 28-JUL-2000; 2000WO-US20710.
 PR 11-AUG-2000; 2000WO-US22031.
 PR 23-AUG-2000; 2000WO-US23228.
 PR 24-AUG-2000; 2000WO-US23328.
 PR 08-NOV-2000; 2000WO-US30952.
 PR 01-DEC-2000; 2000WO-US32678.
 PR 28-FEB-2001; 2001WO-US06520.
 PR 01-JUN-2001; 2001WO-US17800.
 PR 20-JUN-2001; 2001WO-US19692.
 PR 29-JUN-2001; 2001WO-US21066.
 PR 09-JUL-2001; 2001WO-US21735.
 PR 16-JUN-1997; 97US-C49787P.
 PR 12-OCT-1997; 97US-C62250P.
 PR 17-NOV-1997; 97US-C65186P.
 PR 13-NOV-1997; 97US-C65311P.
 PR 25-NOV-1997; 97US-C66770P.
 PR 25-FEB-1998; 98US-C075945P.
 PR 28-MAR-1998; 98US-C078910P.
 PR 28-APR-1998; 98US-C083322P.
 PR 07-MAY-1998; 98US-C084600P.
 PR 28-MAY-1998; 98US-C087106P.
 PR 02-JUN-1998; 98US-C087607P.
 PR 02-JUN-1998; 98US-C087609P.
 PR 02-JUN-1998; 98US-C087759P.
 PR 03-JUN-1998; 98US-C087827P.
 PR 04-JUN-1998; 98US-C088021P.
 PR 04-JUN-1998; 98US-C088025P.
 PR 04-JUN-1998; 98US-C088026P.
 PR 04-JUN-1998; 98US-C088028P.
 PR 04-JUN-1998; 98US-C088029P.
 PR 04-JUN-1998; 98US-C088030P.
 PR 04-JUN-1998; 98US-C088033P.
 PR 04-JUN-1998; 98US-C088036P.
 PR 05-JUN-1998; 98US-C088167P.
 PR 05-JUN-1998; 98US-C088202P.
 PR 05-JUN-1998; 98US-C088212P.
 PR 05-JUN-1998; 98US-C088217P.
 PR 09-JUN-1998; 98US-C088555P.
 PR 10-JUN-1998; 98US-C088734P.
 PR 10-JUN-1998; 98US-C088738P.
 PR 10-JUN-1998; 98US-C088742P.
 PR 10-JUN-1998; 98US-C088810P.
 PR 10-JUN-1998; 98US-C088824P.
 PR 11-JUN-1998; 98US-C088826P.
 PR 11-JUN-1998; 98US-C088858P.
 PR 11-JUN-1998; 98US-C088861P.
 PR 11-JUN-1998; 98US-C088876P.
 PR 12-JUN-1998; 98US-C089105P.
 PR 16-JUN-1998; 98US-C089440P.
 PR 16-JUN-1998; 98US-C089512P.
 PR 16-JUN-1998; 98US-C089514P.
 PR 17-JUN-1998; 98US-C089532P.
 PR 17-JUN-1998; 98US-C089538P.
 PR 17-JUN-1998; 98US-C089598P.
 PR 17-JUN-1998; 98US-C089599P.
 PR 17-JUN-1998; 98US-C089600P.
 PR 17-JUN-1998; 98US-C089653P.
 PR 18-JUN-1998; 98US-C089801P.
 PR 18-JUN-1998; 98US-C089907P.
 PR 18-JUN-1998; 98US-C089908P.
 PR 28-AUG-2001; 2001US-0941992.
 (GETH) GENENTECH INC.
 PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
 PI Ferrara N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ;
 PI Grimaldi JC, Gurney AL, Kijavlin IJ, Napier MA, Pan J, Paoni NP;
 PI Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WJ;

Zhang Z;
 WPI; 2003-247083/24.
 P-PSDB; ABUS9108.
 Novel isolated PRO polypeptides e.g., PRO826, PRO1068, PRO1184, PRO1346 and PRO1375, which stimulate proliferation of stimulated T-lymphocytes are therapeutically useful for enhancing immune response and in cancer treatments
 Claim 2; Fig 152; 648pp; English.
 The invention describes an isolated human PRO polypeptide. The PRO polypeptides are useful in detecting PRO polypeptides in a sample, in linking a bioactive molecule to a cell expressing a PRO polypeptide, and in modulating at least one biological activity of a cell expressing a PRO polypeptide. PRO1312 stimulates hypertrophy of neonatal heart and is thus useful for treating cardiac insufficiency disorders. PRO1154 and PRO1186 stimulate adrenal cortical capillary endothelial growth, and PRO536, PRO943, PRO828, PRO1068 or PRO535, PRO826, PRO819, PRO1126, PRO1360 and PRO1387 induce c-fos in endothelial cells, and are thus useful for treating conditions or disorders where angiogenesis would be beneficial, e.g. wound healing and antagonist of this polypeptide are useful for treating cancerous tumours. PRO812 inhibits vascular endothelial growth factor (VEGF) stimulated proliferation of endothelial cells and is thus useful for inhibiting endothelial cell growth in mammals which would be beneficial in inhibiting tumour growth. PRO826, PRO1068, PRO1184, PRO1346 and PRO1375 stimulate proliferation of stimulated T-lymphocytes and are therapeutically useful for enhancing immune response. PRO826, PRO1068 or PRO1312 enhance survival of retinal neurons cells (PRO1132 is also enhances survival/proliferation of rod photoreceptor cells) and therefore are useful for treating retinal disorders of injuries, e.g. retinitis pigmentosa, AMD. PRO819, PRO813 and PRO1066 induce proliferation of mammalian kidney mesangial cells, and therefore are useful for treating kidney disorders associated with decreased mesangial cell function such as Berger disease or Crohn's nephropathies associated with dermatitis, herpetiformis or Crohn's disease. PRO1310, PRO844, PRO1312, PRO1192 and PRO1387 induce the proliferation and/or redifferentiation of chondrocytes in culture and are thus useful for treating sports injuries, and arthritis. This sequence represents a novel human PRO protein polynucleotide.
 Sequence 992 BP; 235 A; 301 C; 282 G; 174 T; 0 other;

Query Match	100.0%;	Score 992;	DB 25;	Length 992;
Best Local Similarity	100.0%;	Pred. No. 2.7e-236;		
Matches 992;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy 1	GGCACGAGCCAGGAACTAGGAGGTTCTCACTGCCGAGCAGAGGCCCTACACCCACCGAG	60		
Db 1	GGCACGAGCCAGGAACTAGGAGGTTCTCACTGCCGAGCAGAGGCCCTACACCCACCGAG	60		
Qy 61	GCATGGGGCTCCCTGGGCTGTTCCTGTTGGCGGTGGTGGTGGCAGAGCTTCTCCAAAG	120		
Db 61	GCATGGGGCTCCCTGGGCTGTTCCTGTTGGCGGTGGTGGTGGCAGAGCTTCTCCAAAG	120		
Qy 121	CACGGGAGGAGAAATTAACCCCTGCTGCTCCATTTGCTCAAGTCTCGAAGTTTTC	180		
Db 121	CACGGGAGGAGAAATTAACCCCTGCTGCTCCATTTGCTCAAGTCTCGAAGTTTTC	180		
Qy 181	CCAAAGCCCGCTGGGTGCTCATTAACCTGCTGTGCACCCGAGCAGCCACCCATCATCCT	240		
Db 181	CCAAAGCCCGCTGGGTGCTCATTAACCTGCTGTGCACCCGAGCAGCCACCCATCATCCT	240		
Qy 241	ATTCCTCTCTGTGGAA CCAAGAACATCAAGGTGGCCAGAGAGGTGGTGAAGCCACGAGC	300		
Db 241	ATTCCTCTCTGTGGAA CCAAGAACATCAAGGTGGCCAGAGAGGTGGTGAAGCCACGAGC	300		
Qy 301	CGGCTCCTTCAACCTCAAGTCACTCAAGTCCAGTCCAGACCTCTCACCTACTTCT	360		
Db 301	CGGCTCCTTCAACCTCAAGTCACTCAAGTCCAGTCCAGACCTCTCACCTACTTCT	360		
Qy 361	GCCGGGGCTCTCCACCTCCTAGGTGCCCATGTGGACAGTGGCCGAGCTACAGATGACTCTGGG	420		

Db 601 CCAACTTCTCTTCCCTGCGCAGCCAGACATCGGACTGGTTCTGTGTCAGGCTGCAACA 660
QY 661 ACGCCAATGTCCAGCAGCGCCCTCACAGTGGTGCCCGCCAGGTGGTGACCAAGATGG 720
Db 661 ACGCCAATGTCCAGCAGCGCCCTCACAGTGGTGCCCGCCAGGTGGTGACCAAGATGG 720
QY 721 AGCACTGGCAGGTCCTCTGAGAGCCCAATCCTTGGCTTGGCCCTTACAGGACACCC 780
Db 721 AGCACTGGCAGGTCCTCTGAGAGCCCAATCCTTGGCTTGGCCCTTACAGGACACCC 780
QY 781 GCGGTCTGACTGAAGAGGATTTGGGGGTTCCAGATAGGAGTGGGAGGTGAGAGGAC 840
Db 781 GCGGTCTGACTGAAGAGGATTTGGGGGTTCCAGATAGGAGTGGGAGGTGAGAGGAC 840
QY 841 GCAAAGCAGCAGCCATGTAGATGAACCGTCCAGAGAGCCAGCAGCGGACGAGCACTGCA 900
Db 841 GCAAAGCAGCAGCCATGTAGATGAACCGTCCAGAGAGCCAGCAGCGGACGAGCACTGCA 900
QY 901 GGCATCAGGTGCATGTTCTGATTTGGAGTTCATGCAAAATGAGTGTGTTTACTGTC 960
Db 901 GGCATCAGGTGCATGTTCTGATTTGGAGTTCATGCAAAATGAGTGTGTTTACTGTC 960
QY 961 TCTTGCCACAAAAAATAAAAAAAAAAAAAA 992
Db 961 TCTTGCCACAAAAAATAAAAAAAAAAAAAA 992

RESULT 6
ABX81154
ID ABX81154 standard; DNA; 992 BP.
XX AC ABX81154;
XX DT 22-APR-2003 (first entry)
XX DE Novel human secreted or transmembrane protein PRO791 DNA.
XX KW Human; PRO; hypertrophy of neonatal heart; angiogenesis; wound healing;
KW cardiac insufficiency disorder; cancer; tumour; immune response;
KW adrenal cortical capillary endothelial growth; c-fos induction;
KW vascular endothelial growth factor inhibition; VEGF inhibition;
KW endothelial cell growth inhibitor; T-lymphocytes stimulation;
KW retinal neurons cell survival; rod photoreceptor cell survival;
KW retinal disorder; retinitis pigmentosa; kidney disorder;
KW mammalian kidney mesangial cell proliferation; Berger disease;
KW dermatitis; herpeticiformis; Crohn's disease; chondrocyte proliferation;
KW chondrocyte redifferentiation; sports injury; arthritis; gene; ds.
XX OS Homo sapiens.
XX FN US2003027985-A1.
XX PD 06-FEB-2003.
XX PF 14-NOV-2001; 2001US-0990562.
XX QY 05-NOV-1997; 97WO-US20069.
PR 16-SEP-1998; 98WO-US19330.
PR 17-SEP-1998; 98WO-US19437.
PR 07-OCT-1998; 98WO-US21141.
PR 01-DEC-1998; 98WO-US21108.
PR 05-JAN-1999; 99WO-US00106.
PR 08-MAR-1999; 99WO-US05028.
PR 02-JUN-1999; 99WO-US21252.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28301.
PR 01-DEC-1999; 99WO-US28634.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 05-JAN-2000; 2000WO-US00219.

PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US04914.
PR 02-MAR-2000; 2000WO-US05004.
PR 10-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06319.
PR 20-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US07377.
PR 15-MAY-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13358.
PR 22-MAY-2000; 2000WO-US13705.
PR 30-MAY-2000; 2000WO-US14042.
PR 02-JUN-2000; 2000WO-US14941.
PR 28-JUL-2000; 2000WO-US15264.
PR 11-AUG-2000; 2000WO-US20710.
PR 23-AUG-2000; 2000WO-US22031.
PR 24-AUG-2000; 2000WO-US23522.
PR 08-NOV-2000; 2000WO-US23328.
PR 01-DEC-2000; 2000WO-US30952.
PR 28-FEB-2001; 2000WO-US32678.
PR 01-JUN-2001; 2001WO-US06520.
PR 20-JUN-2001; 2001WO-US17800.
PR 29-JUN-2001; 2001WO-US19692.
PR 09-JUL-2001; 2001WO-US21066.
PR 16-JUL-2001; 2001WO-US21735.
PR 17-OCT-1997; 97US-049787P.
PR 12-NOV-1997; 97US-062250P.
PR 13-NOV-1997; 97US-065311P.
PR 24-NOV-1997; 97US-066770P.
PR 25-FEB-1998; 98US-075945P.
PR 20-MAR-1998; 98US-078910P.
PR 28-APR-1998; 98US-083322P.
PR 07-MAY-1998; 98US-084600P.
PR 28-MAY-1998; 98US-087106P.
PR 02-JUN-1998; 98US-087607P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088021P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088026P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088030P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088555P.
PR 10-JUN-1998; 98US-088734P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088742P.
PR 10-JUN-1998; 98US-088810P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088858P.
PR 11-JUN-1998; 98US-088861P.
PR 12-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089440P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089532P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089599P.
PR 17-JUN-1998; 98US-089600P.
PR 17-JUN-1998; 98US-089653P.

```
PR 18-JUN-1998; 98US-089801P;
PR 18-JUN-1998; 98US-089907P;
PR 18-JUN-1998; 98US-089908P;
PR 19-JUN-1998; 98US-089947P;
PR 19-JUN-1998; 98US-089948P;
PR 19-JUN-1998; 98US-089952P;
PR 22-JUN-1998; 98US-090246P;
PR 22-JUN-1998; 98US-090252P;
PR 22-JUN-1998; 98US-090254P;
PR 23-JUN-1998; 98US-090349P;
PR 23-JUN-1998; 98US-090355P;
PR 24-JUN-1998; 98US-090429P;
PR 24-JUN-1998; 98US-090431P;
PR 24-JUN-1998; 98US-090435P;
PR 24-JUN-1998; 98US-090444P;
PR 24-JUN-1998; 98US-090445P;
PR 24-JUN-1998; 98US-090472P;
PR 24-JUN-1998; 98US-090535P;
PR 24-JUN-1998; 98US-090540P;
PR 24-JUN-1998; 98US-090542P;
PR 24-JUN-1998; 98US-090557P;
PR 25-JUN-1998; 98US-090676P;
PR 25-JUN-1998; 98US-090678P;
PR 25-JUN-1998; 98US-090690P;
PR 25-JUN-1998; 98US-090694P;
PR 25-JUN-1998; 98US-090695P;
PR 25-JUN-1998; 98US-090696P;
PR 26-JUN-1998; 98US-090862P;
PR 26-JUN-1998; 98US-090863P;
PR 01-JUL-1998; 98US-091360P;
PR 01-JUL-1998; 98US-091544P;
PR 02-JUL-1998; 98US-091478P;
PR 02-JUL-1998; 98US-091519P;
PR 02-JUL-1998; 98US-091626P;
PR 02-JUL-1998; 98US-091628P;
PR 02-JUL-1998; 98US-091633P;
PR 02-JUL-1998; 98US-091646P;
PR 02-JUL-1998; 98US-091673P;
PR 07-JUL-1998; 98US-091978P;
PR 07-JUL-1998; 98US-091982P;
PR 09-JUL-1998; 98US-092182P;
PR 10-JUL-1998; 98US-092472P;
PR 20-JUL-1998; 98US-093339P;
PR 30-JUL-1998; 98US-094651P;
PR 04-AUG-1998; 98US-095282P;
PR 04-AUG-1998; 98US-095285P;
PR 04-AUG-1998; 98US-095301P;
PR 04-AUG-1998; 98US-095302P;
PR 04-AUG-1998; 98US-095318P;
PR 04-AUG-1998; 98US-095321P;
PR 04-AUG-1998; 98US-095325P;
PR 10-AUG-1998; 98US-095916P;
PR 10-AUG-1998; 98US-095929P;
PR 10-AUG-1998; 98US-096012P;
PR 11-AUG-1998; 98US-096143P;
PR 12-AUG-1998; 98US-096146P;
PR 12-AUG-1998; 98US-096329P;
PR 17-AUG-1998; 98US-096757P;
PR 17-AUG-1998; 98US-096766P;
PR 17-AUG-1998; 98US-096768P;
PR 17-AUG-1998; 98US-096773P;
PR 17-AUG-1998; 98US-096791P;
PR 17-AUG-1998; 98US-096867P;
PR 17-AUG-1998; 98US-096891P;
PR 17-AUG-1998; 98US-096894P;
PR 17-AUG-1998; 98US-096895P;
PR 17-AUG-1998; 98US-096897P;
PR 18-AUG-1998; 98US-096949P;
PR 18-AUG-1998; 98US-096950P;
PR 18-AUG-1998; 98US-096959P;
PR 18-AUG-1998; 98US-096960P;
PR 18-AUG-1998; 98US-097022P;
PR 19-AUG-1998; 98US-097141P;
PR 20-AUG-1998; 98US-097218P;
PR 24-AUG-1998; 98US-097661P;
PR 26-AUG-1998; 98US-097952P;
PR 26-AUG-1998; 98US-097954P;
PR 26-AUG-1998; 98US-097955P;
PR 26-AUG-1998; 98US-097971P;
PR 26-AUG-1998; 98US-097974P;
PR 26-AUG-1998; 98US-097978P;
PR 26-AUG-1998; 98US-097979P;
PR 26-AUG-1998; 98US-097986P;
PR 26-AUG-1998; 98US-098014P;
PR 20-AUG-1998; 98US-097218P;
PR 24-AUG-1998; 98US-097661P;
PR 26-AUG-1998; 98US-097952P;
PR 26-AUG-1998; 98US-097954P;
PR 26-AUG-1998; 98US-097955P;
PR 26-AUG-1998; 98US-097971P;
PR 26-AUG-1998; 98US-097974P;
PR 26-AUG-1998; 98US-097978P;
PR 26-AUG-1998; 98US-097979P;
PR 26-AUG-1998; 98US-097986P;
PR 26-AUG-1998; 98US-098014P;
Query Match 100.0%; Score 992; DB 25; Length 992;
Best Local Similarity 100.0%; Pred. No. 2,7e-236; Indels 0; Gaps 0;
Matches 992; Conservative 0; Mismatches 0;
Qy 1 GGCACGAGCCAGGAACCTAGGAGGTTCTCACTGCCCGAGCAGAGCCCTACACCCACCGAG 60
Db 1 GGCACGAGCCAGGAACCTAGGAGGTTCTCACTGCCCGAGCAGAGCCCTACACCCACCGAG 60
Qy 61 GCATGGGGTCCCTGGGCTGTTCTGCTGGCGTGCTGGCTGCCAGCAGCTTCTCCAAG 120
Db 61 GCATGGGGTCCCTGGGCTGTTCTGCTGGCGTGCTGGCTGCCAGCAGCTTCTCCAAG 120
Qy 121 CACGGGAGGAAGAAATTACCCCTGCTGCTCCATTGCTCAAAAGTCTCGAAGTTTTC 180
Db 121 CACGGGAGGAAGAAATTACCCCTGCTGCTCCATTGCTCAAAAGTCTCGAAGTTTTC 180
Qy 181 CCAAGGCCGCTGGGTGCTCATTAACCTGCTGTGCACCCAGCCACCCACCTACCT 240
Db 181 CCAAGGCCGCTGGGTGCTCATTAACCTGCTGTGCACCCAGCCACCCACCTACCT 240
Qy 241 ATTCCCTCTGTGGAACCAAGAAACATCAAGGTGGCCAAAGAGGTGGTGAAGCCACGAGC 300
Db 241 ATTCCCTCTGTGGAACCAAGAAACATCAAGGTGGCCAAAGAGGTGGTGAAGCCACGAGC 300
Qy 301 CGGCTCCTTCAACCTCAAGTCACTCAAGTCCAGTCCAGTCCAGTCCCTCCTCCTACTTCT 360
Db 301 CGGCTCCTTCAACCTCAAGTCACTCAAGTCCAGTCCAGTCCAGTCCCTCCTCCTACTTCT 360
Qy 361 GCGGGGCTCCTCCACCTCAGGTGCCATGTGGACAGTGCAGGCTACAGATGACTGGG 420
Db 361 GCGGGGCTCCTCCACCTCAGGTGCCATGTGGACAGTGCAGGCTACAGATGACTGGG 420
Qy 421 AGCTGTGTTCAAGCAGTGTCTGAGCTGGGGGCAACTTCACTCTCAGGACAGAGGG 480
Db 421 AGCTGTGTTCAAGCAGTGTCTGAGCTGGGGGCAACTTCACTCTCAGGACAGAGGG 480
Qy 481 CAGSCCCAGGTTGGAGATGATCTGCCAGGCTCCTCGGGCAGCCACCTATCACCACCA 540
Db 481 CAGSCCCAGGTTGGAGATGATCTGCCAGGCTCCTCGGGCAGCCACCTATCACCACCA 540
Qy 541 GCCTGATCGGGAAGGATGGCAGGTCCACTGTCAGCAGAGACCATGTCACAGGAGCCTG 600
Db 541 GCCTGATCGGGAAGGATGGCAGGTCCACTGTCAGCAGAGACCATGTCACAGGAGCCTG 600
Qy 601 CCAACTTCTCCTCTGCGAGCAGACATCGACATGCTGTTCTGGTGCAGGCTGCAACA 660
Db 601 CCAACTTCTCCTCTGCGAGCAGACATCGACATGCTGTTCTGGTGCAGGCTGCAACA 660
Qy 661 ACGCCAATGTCACAGCAGCGCCCTCACAGTGTGTCGCCCCAGGTGGTGAACAGAGATGG 720
Db 661 ACGCCAATGTCACAGCAGCGCCCTCACAGTGTGTCGCCCCAGGTGGTGAACAGAGATGG 720
Qy 721 AGGACTGGCAGGTTCCCTGGAGAGCCCATCCTTGGCTTGGCTTACAGGACACCC 780
Db 721 AGGACTGGCAGGTTCCCTGGAGAGCCCATCCTTGGCTTGGCTTACAGGACACCC 780
Qy 781 GCCGCTCTGAGTGAAGAGGAGTTTGGGGGTTTCAAGATAGGATGGGAGGTGAGAGGAC 840
Db 781 GCCGCTCTGAGTGAAGAGGAGTTTGGGGGTTTCAAGATAGGATGGGAGGTGAGAGGAC 840
Qy 841 GCAAGCAGCAGCCATGTTAGATGAACCGTCCAGAGAGCCAGCAGCGGACGAGGACTGCA 900
```

```

Db      841  GCAAAGCAGCAGCATGTAGTAATGACCGTCCAGAGAGCCAGCAGCGCAGAGACTGCA 900
QY      901  GGCATCAGCGTGCACCTGTTTCGTAATTTGGAGTTTCATGCAAAATGAGTGTGTTTACGTGC 960
Db      901  GGCATCAGCGTGCACCTGTTTCGTAATTTGGAGTTTCATGCAAAATGAGTGTGTTTACGTGC 960
QY      961  TCTTGCACAAAAA 992
Db      961  TCTTGCACAAAAA 992

RESULT 7
ABX90244
ID      ABX90244 standard; cDNA; 992 BP.
XX
AC      ABX90244;
XX
DT      01-MAY-2003 (first entry)
XX
DE      Human secreted/transmembrane protein cDNA, #90.
XX
KW      Human; gene; ss; PRO; secreted; transmembrane; signal peptide;
KW      pharmaceutical; diagnostic; therapeutic; gene therapy.
XX
OS      Homo sapiens.
XX
PN      US2002160384-A1.
XX
PD      31-OCT-2002.
XX
PF      14-NOV-2001; 2001US-0992598.
XX
PR      05-NOV-1997; 97WO-US20069.
PR      16-SEP-1998; 98WO-US19330.
PR      17-SEP-1998; 98WO-US19437.
PR      07-OCT-1998; 98WO-US21141.
PR      01-DEC-1998; 98WO-US25108.
PR      05-JAN-1999; 99WO-US00106.
PR      08-MAR-1999; 99WO-US05028.
PR      02-JUN-1999; 99WO-US12252.
PR      15-SEP-1999; 99WO-US21090.
PR      15-SEP-1999; 99WO-US21547.
PR      30-NOV-1999; 99WO-US28313.
PR      01-DEC-1999; 99WO-US28301.
PR      01-DEC-1999; 99WO-US28634.
PR      16-DEC-1999; 99WO-US30095.
PR      20-DEC-1999; 99WO-US30911.
PR      05-JAN-2000; 2000WO-US00219.
PR      06-JAN-2000; 2000WO-US00376.
PR      11-FEB-2000; 2000WO-US03565.
PR      18-FEB-2000; 2000WO-US04341.
PR      22-FEB-2000; 2000WO-US04414.
PR      24-FEB-2000; 2000WO-US04914.
PR      24-FEB-2000; 2000WO-US05004.
PR      02-MAR-2000; 2000WO-US05841.
PR      10-MAR-2000; 2000WO-US06319.
PR      15-MAR-2000; 2000WO-US06884.
PR      20-MAR-2000; 2000WO-US07377.
PR      30-MAR-2000; 2000WO-US08439.
PR      15-MAY-2000; 2000WO-US13358.
PR      17-MAY-2000; 2000WO-US13705.
PR      22-MAY-2000; 2000WO-US14042.
PR      30-MAY-2000; 2000WO-US14941.
PR      02-JUN-2000; 2000WO-US15264.
PR      28-JUL-2000; 2000WO-US20710.
PR      11-AUG-2000; 2000WO-US22031.
PR      23-AUG-2000; 2000WO-US23522.
PR      24-AUG-2000; 2000WO-US23328.
PR      08-NOV-2000; 2000WO-US30952.
PR      01-DEC-2000; 2000WO-US32678.
PR      28-FEB-2001; 2001WO-US06520.
PR      01-JUN-2001; 2001WO-US17800.

```

```

PR      20-JUN-2001; 2001WO-US19692.
PR      29-JUN-2001; 2001WO-US21066.
PR      09-JUL-2001; 2001WO-US21735.
PR      16-JUN-1997; 97US-049787P.
PR      17-OCT-1997; 97US-062250P.
PR      12-NOV-1997; 97US-065186P.
PR      13-NOV-1997; 97US-065311P.
PR      24-NOV-1997; 97US-066770P.
PR      25-FEB-1998; 98US-075945P.
PR      20-MAR-1998; 98US-078910P.
PR      28-APR-1998; 98US-083322P.
PR      07-MAY-1998; 98US-084600P.
PR      28-MAY-1998; 98US-087106P.
PR      02-JUN-1998; 98US-087607P.
PR      02-JUN-1998; 98US-087609P.
PR      02-JUN-1998; 98US-087759P.
PR      03-JUN-1998; 98US-087827P.
PR      04-JUN-1998; 98US-088021P.
PR      04-JUN-1998; 98US-088025P.
PR      04-JUN-1998; 98US-088026P.
PR      04-JUN-1998; 98US-088028P.
PR      04-JUN-1998; 98US-088029P.
PR      04-JUN-1998; 98US-088030P.
PR      04-JUN-1998; 98US-088033P.
PR      04-JUN-1998; 98US-088326P.
PR      05-JUN-1998; 98US-088467P.
PR      05-JUN-1998; 98US-088202P.
PR      05-JUN-1998; 98US-088212P.
PR      05-JUN-1998; 98US-088217P.
PR      09-JUN-1998; 98US-086555P.
PR      10-JUN-1998; 98US-088734P.
PR      10-JUN-1998; 98US-088738P.
PR      10-JUN-1998; 98US-088742P.
PR      10-JUN-1998; 98US-088810P.
PR      10-JUN-1998; 98US-088824P.
PR      10-JUN-1998; 98US-088826P.
PR      11-JUN-1998; 98US-088858P.
PR      11-JUN-1998; 98US-088861P.
PR      11-JUN-1998; 98US-088876P.
PR      12-JUN-1998; 98US-089105P.
PR      16-JUN-1998; 98US-089440P.
PR      16-JUN-1998; 98US-089512P.
PR      16-JUN-1998; 98US-089514P.
PR      17-JUN-1998; 98US-089532P.
PR      17-JUN-1998; 98US-089538P.
PR      17-JUN-1998; 98US-089598P.
PR      17-JUN-1998; 98US-089599P.
PR      17-JUN-1998; 98US-089600P.
PR      17-JUN-1998; 98US-089653P.
PR      18-JUN-1998; 98US-089801P.
PR      18-JUN-1998; 98US-089907P.
PR      18-JUN-1998; 98US-089908P.
PR      28-AUG-2001; 2001US-0941992.
XX
XX      (GETH ) GENENTECH INC.
XX
PI      Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
PI      Ferrara N, Fong S, Gerber H, Gerritsen WB, Goddard A, Godowski PJ;
PI      Grimaldi JC, Gurney AL, Kljavin IJ, Napier MA, Pan J, Paoni NF;
PI      Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI;
PI      Zhang Z;
XX
XX      WPI; 2003-288106/28.
DR      P-PSDB; AB060538.
XX
PT      New transmembrane polypeptides and nucleic acids encoding the
PT      polypeptides, useful in gene therapy, in chromosome identification, as
PT      chromosome markers, or in generating probes -
XX
PS      Claim 2; Fig 150; 650pp; English.
XX
CC      The invention discloses isolated PRO secreted/transmembrane polypeptides
CC      comprising a sequence without signal peptide and the nucleic acid

```


CC encoding them. The polypeptides can be used to raise antibodies that
CC specifically bind to the PRO polypeptide, for linking a bioactive
CC molecule to a cell expressing a PRO protein and for modulating at least
CC one biological activity of a cell. The PRO polypeptides or
CC polynucleotides are also useful in gene therapy, in chromosome
CC identification, as chromosome markers, or in generating probes. The PRO
CC polypeptides are useful as molecular markers for protein
CC electrophoresis, and the isolated nucleic acids may be used for
CC recombinantly expressing those markers. The PRO polypeptides and nucleic
CC acids may also be used in tissue typing. Anti-PRO antibodies are useful
CC in diagnostic assays for PRO, and in affinity purification of PRO from
CC recombinant cell culture or natural sources. The sequences presented in
CC ABX90083-ABX90468 are the genes encoding, the primers amplifying and the
CC probes detecting the PRO polynucleotides of the invention.
CC Note: The sequence data for this patent is also available in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX

SQ Sequence 992 BP; 235 A; 301 C; 282 G; 174 T; 0 other;

Query Match 100.0%; Score 992; DB 25; Length 992;
Best Local Similarity 100.0%; Pred. No. 2.7e-236;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GGCAGAGCCAGGACTAGGAGTTCTCACTGCCGAGCAGAGCCCTACACCCACCGAG	60
Db	1	GGCAGAGCCAGGAACTAGGAGTTCTCACTGCCGAGCAGAGCCCTACACCCACCGAG	60
QY	61	GCATGGGGCTCCCTGGGCTGTCTGCTGGCGCTGCTGCCAGCAGCTTCTCCAAGG	120
Db	61	GCATGGGGCTCCCTGGGCTGTCTGCTGGCGCTGCTGCCAGCAGCTTCTCCAAGG	120
QY	121	CACGGAGAGAAATACCCCTGTGTTCTCATTCGCTCAAAAGTCTCGAAGTTTTC	180
Db	121	CACGGAGAGAAATACCCCTGTGTTCTCATTCGCTCAAAAGTCTCGAAGTTTTC	180
QY	181	CCAAAGGCGCTGGTGTCTATACCTCTGTGCACCCGACCCACCCGCTCATCCCT	240
Db	181	CCAAAGGCGCTGGTGTCTATACCTCTGTGCACCCGACCCACCCGCTCATCCCT	240
QY	241	ATTCCTCTGTGGAAACCAAGAACATCAAGTGGCCAGAGGTGGTGAAGACCCACGAGC	300
Db	241	ATTCCTCTGTGGAAACCAAGAACATCAAGTGGCCAGAGGTGGTGAAGACCCACGAGC	300
QY	301	CGGCTCTCTCAACTCAAGTCACTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT	360
Db	301	CGGCTCTCTCAACTCAAGTCACTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT	360
QY	361	CGCGGGCTCTCCACCTCAGTGCCTATGAGTGGCAGTGGCAGGCTACAGTGCACCTGGG	420
Db	361	CGCGGGCTCTCCACCTCAGTGCCTATGAGTGGCAGTGGCAGGCTACAGTGCACCTGGG	420
QY	421	AGCTGTGTCCAAAGCAGTGTCTGAGTGTGGGCAACTTCACTCTGCAGGACAGAGGGG	480
Db	421	AGCTGTGTCCAAAGCAGTGTCTGAGTGTGGGCAACTTCACTCTGCAGGACAGAGGGG	480
QY	481	CAGGCCCGAGGTGGAGATGATCGCAGGGCTCTGGGCGAGCCACCTATCAACA	540
Db	481	CAGGCCCGAGGTGGAGATGATCGCAGGGCTCTGGGCGAGCCACCTATCAACA	540
QY	541	GCCTGATCGGAGGATGGGCGAGTGCACCTGCAGCAGAGACCATGCCACAGGCGCTG	600
Db	541	GCCTGATCGGAGGATGGGCGAGTGCACCTGCAGCAGAGACCATGCCACAGGCGCTG	600
QY	601	CCAACTTCTCTCTCTGCGGAGCAGACATCGGACTGGTCTGGTGGCCAGGCTCAACA	660
Db	601	CCAACTTCTCTCTCTGCGGAGCAGACATCGGACTGGTCTGGTGGCCAGGCTCAACA	660
QY	661	ACGCCAATGTCCAGCAGAGCCCTCAGTGTGGTGGCCCGGAGTGGTGAACAGAGATGG	720
Db	661	ACGCCAATGTCCAGCAGAGCCCTCAGTGTGGTGGCCCGGAGTGGTGAACAGAGATGG	720
QY	721	AGGACTGGCAGGGTCCCTCTGGAGAGCCCATCCCTTGGCTTCCGCTCTACAGGAGCACC	780
Db	721	AGGACTGGCAGGGTCCCTCTGGAGAGCCCATCCCTTGGCTTCCGCTCTACAGGAGCACC	780

Db	721	AGGACTGGCAGGGTCCCTCTGGAGAGCCCATCCCTTGGCTTCCGCTCTACAGGAGCACC	780
QY	781	GCCGTCTCAGTGAAGAGGAGTTTCGGGGGTTTCAGATAGGGAATGGGGAGGTTCAGAGGAC	840
Db	781	GCCGTCTCAGTGAAGAGGAGTTTCGGGGGTTTCAGATAGGGAATGGGGAGGTTCAGAGGAC	840
QY	841	GCAAGCAGCAGCCCATGTAGAAATGAACCGTCCAGAGAGCCCAAGCAGCAGGAGACTGCA	900
Db	841	GCAAGCAGCAGCCCATGTAGAAATGAACCGTCCAGAGAGCCCAAGCAGCAGGAGACTGCA	900
QY	901	GGCCATCAGCGTGCACCTGTTCTGTTATTTGGAGTTTCATGCAAAATGAGTGTGTTTAGCTGC	960
Db	901	GGCCATCAGCGTGCACCTGTTCTGTTATTTGGAGTTTCATGCAAAATGAGTGTGTTTAGCTGC	960
QY	961	TCTTGGCCACAAAAAATAAAAAAAAAAAAAAAAAAAAAA	992
Db	961	TCTTGGCCACAAAAAATAAAAAAAAAAAAAAAAAAAAAA	992

RESULT 8
ABX77855

ID ABX77855 standard; cDNA; 992 BP.

XX ABX77855;

XX 14-APR-2003 (first entry)

XX Human PRO polynucleotide #61.

XX Human; PRO; gene; ss; cytostatic; tumour; cancer; breast; lung; stomach;
XX liver; horse; cow; dog; cat; sheep; pig; goat; rabbit; ADEPT;
XX antibody-dependent enzyme mediated prodrug therapy.

XX Homo sapiens.

XX US2003027163-A1.

XX 06-FEB-2003.

XX 15-NOV-2001; 2001US-0997666.

XX 05-NOV-1997; 97NO-US20069.

XX 16-SEP-1998; 98WO-US19330.

XX 17-SEP-1998; 98WO-US19437.

XX 07-OCT-1998; 98WO-US21141.

XX 01-DEC-1998; 98WO-US25108.

XX 05-JAN-1999; 99WO-US00106.

XX 08-MAR-1999; 99WO-US05028.

XX 02-JUN-1999; 99WO-US12252.

XX 15-SEP-1999; 99WO-US21090.

XX 30-NOV-1999; 99WO-US28313.

XX 01-DEC-1999; 99WO-US28301.

XX 16-DEC-1999; 99WO-US28634.

XX 20-DEC-1999; 99WO-US30095.

XX 05-JAN-2000; 2000WO-US00219.

XX 06-JAN-2000; 2000WO-US00376.

XX 11-FEB-2000; 2000WO-US03565.

XX 18-FEB-2000; 2000WO-US04341.

XX 24-FEB-2000; 2000WO-US04414.

XX 02-MAR-2000; 2000WO-US05004.

XX 10-MAR-2000; 2000WO-US05841.

XX 15-MAR-2000; 2000WO-US06884.

XX 20-MAR-2000; 2000WO-US07377.

XX 30-MAR-2000; 2000WO-US08439.

XX 15-MAY-2000; 2000WO-US13358.

XX 17-MAY-2000; 2000WO-US13705.

XX 22-MAY-2000; 2000WO-US14042.

XX 30-MAY-2000; 2000WO-US14941.

XX 02-JUN-2000; 2000WO-US15264.

PR 28-JUL-2000; 2000WO-US20710.
PR 11-AUG-2000; 2000WO-US22031.
PR 23-AUG-2000; 2000WO-US23352.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 16-JUN-1997; 97US-049787P.
PR 17-OCT-1997; 97US-062250P.
PR 12-NOV-1997; 97US-065186P.
PR 13-NOV-1997; 97US-065311P.
PR 24-NOV-1997; 97US-066770P.
PR 25-FEB-1998; 98US-075945P.
PR 20-MAR-1998; 98US-078910P.
PR 28-APR-1998; 98US-083322P.
PR 07-MAY-1998; 98US-084600P.
PR 28-MAY-1998; 98US-087106P.
PR 02-JUN-1998; 98US-087607P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088021P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088026P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088030P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088036P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.
PR 10-JUN-1998; 98US-088734P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088742P.
PR 10-JUN-1998; 98US-088810P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088858P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089440P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089532P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089599P.
PR 17-JUN-1998; 98US-089600P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089801P.
PR 18-JUN-1998; 98US-089907P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089947P.
PR 19-JUN-1998; 98US-089948P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 23-JUN-1998; 98US-090349P.
PR 23-JUN-1998; 98US-090355P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090431P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090445P.
PR 24-JUN-1998; 98US-090472P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 24-JUN-1998; 98US-090542P.
PR 24-JUN-1998; 98US-090557P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 01-JUL-1998; 98US-091360P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091519P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091633P.
PR 02-JUL-1998; 98US-091646P.
PR 02-JUL-1998; 98US-091673P.
PR 07-JUL-1998; 98US-091978P.
PR 09-JUL-1998; 98US-092182P.
PR 10-JUL-1998; 98US-092472P.
PR 20-JUL-1998; 98US-093339P.
PR 30-JUL-1998; 98US-094651P.
PR 04-AUG-1998; 98US-095282P.
PR 04-AUG-1998; 98US-095285P.
PR 04-AUG-1998; 98US-095301P.
PR 04-AUG-1998; 98US-095302P.
PR 04-AUG-1998; 98US-095318P.
PR 04-AUG-1998; 98US-095321P.
PR 04-AUG-1998; 98US-095325P.
PR 10-AUG-1998; 98US-095916P.
PR 10-AUG-1998; 98US-095929P.
PR 10-AUG-1998; 98US-096012P.
PR 11-AUG-1998; 98US-096143P.
PR 11-AUG-1998; 98US-096146P.
PR 12-AUG-1998; 98US-096329P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096768P.
PR 17-AUG-1998; 98US-096773P.
PR 17-AUG-1998; 98US-096791P.
PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096891P.
PR 17-AUG-1998; 98US-096894P.
PR 17-AUG-1998; 98US-096895P.
PR 18-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-096950P.
PR 18-AUG-1998; 98US-096959P.
PR 18-AUG-1998; 98US-096960P.
PR 18-AUG-1998; 98US-097022P.
PR 19-AUG-1998; 98US-097141P.
PR 20-AUG-1998; 98US-097218P.
PR 24-AUG-1998; 98US-097661P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.
PR 26-AUG-1998; 98US-097978P.
PR 26-AUG-1998; 98US-097979P.
PR 26-AUG-1998; 98US-097986P.
PR 31-AUG-1998; 98US-098014P.
PR 31-AUG-1998; 98US-098525P.
PR 16-SEP-1998; 98US-100634P.
PR 17-SEP-1998; 98US-100858P.
PR 22-DEC-1998; 98US-113296P.
PR 12-MAR-1999; 99US-123957P.

PR 23-JUN-1999; 99US-141037P.
PR 07-JUL-1999; 99US-143048P.

```
Query Match      100.0%; Score 992; DB 25; Length 992;
Best Local Similarity 100.0%; Pred. No. 2.7e-236;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACGAGCCAGAACTAGAGGTTCTACTGCCGAGCAGAGGCGCTACACCCACCCAG 60
Db 1 GGCACGAGCCAGAACTAGAGGTTCTACTGCCGAGCAGAGGCGCTACACCCACCCAG 60

QY 61 GCATGGGGTCCCTGGGCTGTTCTGCTGGCGTGCTGCTGCCAGAGCTTCTCCAAAG 120
Db 61 GCATGGGGTCCCTGGGCTGTTCTGCTGGCGTGCTGCTGCCAGAGCTTCTCCAAAG 120

QY 121 CACGGGAGGAGAAATACCCCTGGTCTCAATGCTTCAAGTCTCGAAGTTTCC 180
Db 121 CACGGGAGGAGAAATACCCCTGGTCTCAATGCTTCAAGTCTCGAAGTTTCC 180

QY 181 CCAAGGCGCTGGGTGCTCATAACTGCTGTGCAACCCAGCCACCGCCCATCACT 240
Db 181 CCAAGGCGCTGGGTGCTCATAACTGCTGTGCAACCCAGCCACCGCCCATCACT 240

QY 241 ATTCCCTCTGTGGAACCAAGAACATCAAGGTGGCCAGAGGTGTGAAGCCACGAGC 300
Db 241 ATTCCCTCTGTGGAACCAAGAACATCAAGGTGGCCAGAGGTGTGAAGCCACGAGC 300

QY 301 CGGCTCCTTCAACCTCAAGCTCACACTCAAGTCAAGTCCAGACTGCTCACTTCT 360
Db 301 CGGCTCCTTCAACCTCAAGCTCACACTCAAGTCCAGACTGCTCACTTCT 360

QY 361 GCCGGGCTCTCCACCTCAGTGTGCCATGTGGACAGTGCAGGCTACAGATGCACCTGG 420
Db 361 GCCGGGCTCTCCACCTCAGTGTGCCATGTGGACAGTGCAGGCTACAGATGCACCTGG 420

QY 421 AGCTGTGTGTCACAGCAAGTGTGAGTGTGGCGGCACTTCACTCTGCAGCAGAGGG 480
Db 421 AGCTGTGTGTCACAGCAAGTGTGAGTGTGGCGGCACTTCACTCTGCAGCAGAGGG 480

QY 481 CAGGCCCGAGGTGAGATGATCTGCCAGGCTCTCGGGCAGCCACCTATCAACCA 540
Db 481 CAGGCCCGAGGTGAGATGATCTGCCAGGCTCTCGGGCAGCCACCTATCAACCA 540

QY 541 GCCTGATCGGAAGGATGGGAGGTCCACCTGCAGCAGAGACCATGCCACAGGAGCTG 600
Db 541 GCCTGATCGGAAGGATGGGAGGTCCACCTGCAGCAGAGACCATGCCACAGGAGCTG 600

QY 601 CCAACTTCTCCTCTGCGGAGCCAGACATCGAGTGGTCTGCGGAGGCTGCAAA 660
Db 601 CCAACTTCTCCTCTGCGGAGCCAGACATCGAGTGGTCTGCGGAGGCTGCAAA 660

QY 661 ACGCAATGTCACAGCCCTCTCAGTGTGGCCCGGAGTGTGACAGAGATGG 720
Db 661 ACGCAATGTCACAGCCCTCTCAGTGTGGCCCGGAGTGTGACAGAGATGG 720

QY 721 AGGACTGGCAGGGTCCCTGAGAGCCCTTCTGCTTGCCTCTTACAGGAGCACC 780
Db 721 AGGACTGGCAGGGTCCCTGAGAGCCCTTCTGCTTGCCTCTTACAGGAGCACC 780

QY 781 GCCCTCTGAGTGAAGAGAGTTTGGGGGTTTCAGATAGGAAATGGGAGGTTCAGAG 840
Db 781 GCCCTCTGAGTGAAGAGAGTTTGGGGGTTTCAGATAGGAAATGGGAGGTTCAGAG 840

QY 841 GCAAGCAGCAGCCATGTAGAAATCAACCTCCAGAGCCAGCCACCGCAGAGACTGCA 900
Db 841 GCAAGCAGCAGCCATGTAGAAATCAACCTCCAGAGCCAGCCACCGCAGAGACTGCA 900

QY 901 GGCCATCAGCGTGCCTGTTCTGATTTGGAGTTCATGCAAAATGAGTGTGTTTGTGTC 960
Db 901 GGCCATCAGCGTGCCTGTTCTGATTTGGAGTTCATGCAAAATGAGTGTGTTTGTGTC 960

QY 961 TCTTGCCACAAAAA 992
Db 961 TCTTGCCACAAAAA 992
```

Db 961 TCTTGCCACAAAAA 992

RESULT 9

```
ABX79451
ID ABX79451 standard; cDNA; 992 BP.
XX AC ABX79451;
XX DT 17-APR-2003 (first entry)
XX XX
XX Human secreted/transmembrane protein cDNA, #90.
XX
XX Human; gene; ss; PRO; secreted; transmembrane; signal peptide;
XX pharmaceutical; diagnostic; biosensor; bioreactor; tumour; therapeutic;
XX colon cancer; lung cancer; breast cancer; cancer; gene therapy.
XX Homo sapiens.
XX
XX US2002142961-A1.
XX
XX 03-OCT-2002.
XX
XX 19-NOV-2001; 2001US-0989721.
XX
XX 05-NOV-1997; 97WO-US20069.
XX 17-SEP-1998; 98WO-US19437.
XX 07-OCT-1998; 98WO-US21141.
XX 01-DEC-1998; 98WO-US25108.
XX 05-JAN-1999; 99WO-US00106.
XX 08-MAR-1999; 99WO-US05028.
XX 02-JUN-1999; 99WO-US12252.
XX 15-SEP-1999; 99WO-US21090.
XX 15-SEP-1999; 99WO-US21547.
XX 30-NOV-1999; 99WO-US28313.
XX 01-DEC-1999; 99WO-US28301.
XX 01-DEC-1999; 99WO-US28634.
XX 16-DEC-1999; 99WO-US30095.
XX 20-DEC-1999; 99WO-US30911.
XX 05-JAN-2000; 2000WO-US00219.
XX 06-JAN-2000; 2000WO-US00376.
XX 11-FEB-2000; 2000WO-US03565.
XX 18-FEB-2000; 2000WO-US04341.
XX 22-FEB-2000; 2000WO-US04414.
XX 24-FEB-2000; 2000WO-US04914.
XX 24-FEB-2000; 2000WO-US05004.
XX 02-MAR-2000; 2000WO-US05841.
XX 10-MAR-2000; 2000WO-US06319.
XX 15-MAR-2000; 2000WO-US06884.
XX 20-MAR-2000; 2000WO-US07377.
XX 30-MAR-2000; 2000WO-US08439.
XX 15-MAY-2000; 2000WO-US13358.
XX 17-MAY-2000; 2000WO-US13705.
XX 22-MAY-2000; 2000WO-US14042.
XX 30-MAY-2000; 2000WO-US14941.
XX 02-JUN-2000; 2000WO-US15264.
XX 28-JUL-2000; 2000WO-US20710.
XX 11-AUG-2000; 2000WO-US22031.
XX 23-AUG-2000; 2000WO-US23522.
XX 24-AUG-2000; 2000WO-US23328.
XX 08-NOV-2000; 2000WO-US30952.
XX 01-DEC-2000; 2000WO-US32678.
XX 28-FEB-2001; 2001WO-US06520.
XX 01-JUN-2001; 2001WO-US17800.
XX 20-JUN-2001; 2001WO-US19692.
XX 29-JUN-2001; 2001WO-US21066.
XX 09-JUL-2001; 2001WO-US21735.
XX 15-JUN-1997; 97US-049787P.
XX 17-OCT-1997; 97US-062250P.
XX 12-NOV-1997; 97US-065186P.
XX 13-NOV-1997; 97US-065311P.
XX 24-NOV-1997; 97US-066770P.
XX 25-FEB-1998; 98US-075945P.
```

PR 20-MAR-1998; 98US-078910P.
 PR 28-APR-1998; 98US-083322P.
 PR 07-MAY-1998; 98US-084600P.
 PR 28-MAY-1998; 98US-087106P.
 PR 02-JUN-1998; 98US-087607P.
 PR 02-JUN-1998; 98US-087609P.
 PR 02-JUN-1998; 98US-087759P.
 PR 03-JUN-1998; 98US-087827P.
 PR 04-JUN-1998; 98US-088021P.
 PR 04-JUN-1998; 98US-088025P.
 PR 04-JUN-1998; 98US-088026P.
 PR 04-JUN-1998; 98US-088028P.
 PR 04-JUN-1998; 98US-088029P.
 PR 04-JUN-1998; 98US-088030P.
 PR 04-JUN-1998; 98US-088033P.
 PR 04-JUN-1998; 98US-088326P.
 PR 05-JUN-1998; 98US-088167P.
 PR 05-JUN-1998; 98US-088202P.
 PR 05-JUN-1998; 98US-088212P.
 PR 05-JUN-1998; 98US-088217P.
 PR 09-JUN-1998; 98US-088655P.
 PR 10-JUN-1998; 98US-088734P.
 PR 10-JUN-1998; 98US-088738P.
 PR 10-JUN-1998; 98US-088742P.
 PR 10-JUN-1998; 98US-088810P.
 PR 10-JUN-1998; 98US-088824P.
 PR 10-JUN-1998; 98US-088826P.
 PR 11-JUN-1998; 98US-088858P.
 PR 11-JUN-1998; 98US-088861P.
 PR 11-JUN-1998; 98US-088876P.
 PR 12-JUN-1998; 98US-089105P.
 PR 16-JUN-1998; 98US-089440P.
 PR 16-JUN-1998; 98US-089512P.
 PR 16-JUN-1998; 98US-089514P.
 PR 17-JUN-1998; 98US-089532P.
 PR 17-JUN-1998; 98US-089538P.
 PR 17-JUN-1998; 98US-089598P.
 PR 17-JUN-1998; 98US-089599P.
 PR 17-JUN-1998; 98US-089600P.
 PR 18-JUN-1998; 98US-089653P.
 PR 18-JUN-1998; 98US-089801P.
 PR 18-JUN-1998; 98US-089907P.
 PR 18-JUN-1998; 98US-089908P.
 PR 28-AUG-2001; 2001US-0941992.

(GETH) GENENTECH INC.

PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL,
 PI Ferrara N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ,
 PI Grimaldi JC, Gurney AL, Kijavini IJ, Napier MA, Pan J, Paoni NF,
 PI Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams FM, Wood WI,
 PI Zhang Z;

XX WPI; 2003-155950/15.
 DR P-PSDB; ABUS8960.

XX New secreted and transmembrane PRO polypeptides (e.g. PRO183, PRO184,
 PT PRO361 or PRO846) useful as targets for therapeutic intervention in
 PT cancers (e.g. lung or breast cancers), or for diagnosing these cancers
 PT -

XX Claim 2; Fig 150; 647pp; English.

XX The invention discloses isolated PRO secreted/transmembrane polypeptides
 CC comprising a sequence without signal peptide and the nucleic acid
 CC encoding them. The polypeptides can be used to raise antibodies that
 CC specifically bind to the PRO polypeptide, for linking a bioactive
 CC molecule to a cell expressing a PRO protein and for modulating at least
 CC one biological activity of a cell. The PRO polypeptides or
 CC polynucleotides are also useful as pharmaceuticals, diagnostics,
 CC biosensors or bioreactors, for detecting or treating e.g. tumours in
 CC mammals, e.g. humans, dogs, cats, cattle, horses, sheep, pigs, goats or
 CC rabbits as targets for therapeutic intervention in certain cancers (e.g.

CC colon, lung or breast cancers) and diagnostic determination of the
 CC presence of these cancers. The PRO polypeptides are also useful as
 CC molecular weight markers or for chromosome identification. The PRO genes
 CC are useful as hybridisation probes or for screening libraries of human
 CC cDNA, genomic DNA or mRNA. The PRO genes may also be used in gene
 CC therapy, particularly for replacing a defective gene. The sequences
 CC presented in ABX79290-ABX79675 are the genes encoding, the primers
 CC amplifying and the probes detecting the PRO polynucleotides of the
 CC invention.
 CC Note: The sequence data for this patent is also available in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 992 BP; 235 A; 301 C; 282 G; 174 T; 0 other;

Query Match 100.0%; Score 992; DB 25; Length 992;
 Best Local Similarity 100.0%; Pred. No. 2,7e-236;
 Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACGAGCCAGGAAC TAGAGGTTCTCACTGCCGAGCAGAGGCCCTACCCACCGAG 60
 Db 1 GGCACGAGCCAGGAAC TAGAGGTTCTCACTGCCGAGCAGAGGCCCTACCCACCGAG 60
 QY 61 GCATGGGGCTCCCTGGGCTGTTCTGCTGGCCGCTGTGCTGCCAGAGTTTCCAAG 120
 Db 61 GCATGGGGCTCCCTGGGCTGTTCTGCTGGCCGCTGTGCTGCCAGAGTTTCCAAG 120
 QY 121 CACGGGAGGAAGAAATTACCCCTGTGTTCTCAATTGCCCTACAAAGTCCTGGAAGTTTTC 180
 Db 121 CACGGGAGGAAGAAATTACCCCTGTGTTCTCAATTGCCCTACAAAGTCCTGGAAGTTTTC 180
 QY 181 CCAAGGCCGCTGGGTGCTCATAAAGTGTGTCACCCAGCCACCCAGCCACCTCACCT 240
 Db 181 CCAAGGCCGCTGGGTGCTCATAAAGTGTGTCACCCAGCCACCCAGCCACCTCACCT 240
 QY 241 ATTCCCTCTGTGGAACCAAGACATCAAGGTGGCCAGAGGTGTGAAGACCCAGAGC 300
 Db 241 ATTCCCTCTGTGGAACCAAGACATCAAGGTGGCCAGAGGTGTGAAGACCCAGAGC 300
 QY 301 CGGCTCTCTTCAACCTCAACGTCACTCAAGTCCAGTCCAGACCTGTCACTACTTCT 360
 Db 301 CGGCTCTCTTCAACCTCAACGTCACTCAAGTCCAGTCCAGACCTGTCACTACTTCT 360
 QY 361 GCGGGGCTCTCAACCTCAACGTCACTCAAGTCCAGTCCAGTCCAGTCCAGTCCAGT 420
 Db 361 GCGGGGCTCTCAACCTCAACGTCACTCAAGTCCAGTCCAGTCCAGTCCAGTCCAGT 420
 QY 421 AGCTGTGTCCAAAGCCAGTGTCTGAGCTGCGGGCCAACTTCACTCTGAGGACAGGGG 480
 Db 421 AGCTGTGTCCAAAGCCAGTGTCTGAGCTGCGGGCCAACTTCACTCTGAGGACAGGGG 480
 QY 481 CAGGCCCCAGGGTGGAGATGATCTGCCAGGGGTCTCTCGGGGAGGCCCACTATCAACA 540
 Db 481 CAGGCCCCAGGGTGGAGATGATCTGCCAGGGGTCTCTCGGGGAGGCCCACTATCAACA 540
 QY 541 GCCTGATCGGGAAGATGGGAGGTCCACCTGTCAGCAGACAGACCATGCCAGGAGCCTG 600
 Db 541 GCCTGATCGGGAAGATGGGAGGTCCACCTGTCAGCAGACAGACCATGCCAGGAGCCTG 600
 QY 601 CCAACTTCTCTTCTCTGCGGAGCCAGACATCGGACTGTTCTTGGTGGCAGGCTGCAACA 660
 Db 601 CCAACTTCTCTTCTCTGCGGAGCCAGACATCGGACTGTTCTTGGTGGCAGGCTGCAACA 660
 QY 661 AGCCCAAGTCCAGCAGACCGCCCTCAAGTGTGTGCCCGAGGTGGTACCAGAGATGG 720
 Db 661 AGCCCAAGTCCAGCAGACCGCCCTCAAGTGTGTGCCCGAGGTGGTACCAGAGATGG 720
 QY 721 AGGACTGCGAGGTTCCTCTGAGAGCCCATCTTTCCTTGGCTTCCCGCTTACAGGAGCACCC 780
 Db 721 AGGACTGCGAGGTTCCTCTGAGAGCCCATCTTTCCTTGGCTTCCCGCTTACAGGAGCACCC 780
 QY 781 GCGTCTGTAGTGAAGAGGAGTTTGGGGGGTTCAGGATAGGGAATGGGAGGTTCAGAGGAC 840
 Db 781 GCGTCTGTAGTGAAGAGGAGTTTGGGGGGTTCAGGATAGGGAATGGGAGGTTCAGAGGAC 840

QY 841 GCAAGCAGCAGCCATGTAGTAATGAACCTCCAGAGAGCCAGCAGCGGAGGACTGCA 900
|||
Db 841 GCAAGCAGCAGCCATGTAGTAATGAACCTCCAGAGAGCCAGCAGCGGAGGACTGCA 900
QY 901 GGCCTCAGCGTGCATGTTCTGATTTGAGTTCATGCAAAATGAGTGTGTTTTAGCTGC 960
|||
Db 901 GGCCTCAGCGTGCATGTTCTGATTTGAGTTCATGCAAAATGAGTGTGTTTTAGCTGC 960
QY 961 TCTTCCACAAAAAATAAAAAAAAAAAAAAAAAA 992
|||
Db 961 TCTTCCACAAAAAATAAAAAAAAAAAAAAAAAA 992

RESULT 10

ABX64090

ID ABX64090 standard; cDNA; 992 BP.

XX AC ABX64090;

XX DT 26-FEB-2003 (first entry)

XX DE cDNA encoding human PRO809 polypeptide.

XX KW Human; PRO polypeptide; secreted protein; transmembrane protein;
XX genetic disorder; antibacterial; immunosuppressive; transgenic;
KW gene therapy; gene; ss.

XX OS Homo sapiens.

XX PN US2002103125-A1.

XX PD 01-AUG-2002.

XX PF 20-NOV-2001; 2001US-0989731.

PR 05-NOV-1997; 97WO-US20069.

PR 16-SEP-1998; 98WO-US19330.

PR 17-SEP-1998; 98WO-US19437.

PR 07-OCT-1998; 98WO-US21141.

PR 01-DEC-1998; 98WO-US25108.

PR 05-JAN-1999; 99WO-US00106.

PR 08-MAR-1999; 99WO-US05028.

PR 02-JUN-1999; 99WO-US12252.

PR 15-SEP-1999; 99WO-US21090.

PR 15-SEP-1999; 99WO-US21547.

PR 30-NOV-1999; 99WO-US28313.

PR 01-DEC-1999; 99WO-US28301.

PR 01-DEC-1999; 99WO-US28634.

PR 16-DEC-1999; 99WO-US30095.

PR 20-DEC-1999; 99WO-US30911.

PR 06-JAN-2000; 2000WO-US00219.

PR 06-JAN-2000; 2000WO-US00376.

PR 11-FEB-2000; 2000WO-US03565.

PR 18-FEB-2000; 2000WO-US04341.

PR 22-FEB-2000; 2000WO-US04414.

PR 24-FEB-2000; 2000WO-US04914.

PR 24-FEB-2000; 2000WO-US05004.

PR 02-MAR-2000; 2000WO-US05841.

PR 10-MAR-2000; 2000WO-US06319.

PR 15-MAR-2000; 2000WO-US06884.

PR 20-MAR-2000; 2000WO-US07377.

PR 30-MAR-2000; 2000WO-US08439.

PR 15-MAY-2000; 2000WO-US13358.

PR 17-MAY-2000; 2000WO-US13705.

PR 22-MAY-2000; 2000WO-US14042.

PR 30-MAY-2000; 2000WO-US14941.

PR 02-JUN-2000; 2000WO-US15264.

PR 28-JUL-2000; 2000WO-US20710.

PR 11-AUG-2000; 2000WO-US22031.

PR 23-AUG-2000; 2000WO-US23522.

PR 24-AUG-2000; 2000WO-US23328.

PR 08-NOV-2000; 2000WO-US30952.

PR 01-DEC-2000; 2000WO-US32678.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 16-JUN-1997; 97US-049787P.
PR 17-OCT-1997; 97US-062250P.
PR 12-NOV-1997; 97US-065186P.
PR 13-NOV-1997; 97US-065311P.
PR 24-NOV-1997; 97US-066770P.
PR 25-FEB-1998; 98US-075945P.
PR 20-MAR-1998; 98US-078910P.
PR 28-APR-1998; 98US-083322P.
PR 07-MAY-1998; 98US-084600P.
PR 28-MAY-1998; 98US-087106P.
PR 02-JUN-1998; 98US-087607P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088021P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088026P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088030P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.
PR 10-JUN-1998; 98US-088734P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088742P.
PR 10-JUN-1998; 98US-088810P.
PR 10-JUN-1998; 98US-088824P.
PR 11-JUN-1998; 98US-088858P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089440P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089532P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089599P.
PR 17-JUN-1998; 98US-089600P.
PR 18-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089801P.
PR 18-JUN-1998; 98US-089907P.
PR 18-JUN-1998; 98US-089908P.
PR 28-AUG-2001; 2001US-0941992.
XX XX
PA (GETH) GENENTECH LTD.
XX XX
PI Askenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
PI Ferrara N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski P;
PI Grimaldi JC, Gurney AL, Kijavini IO, Napier MA, Pan J, Paoni NF;
PI Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI;
PI Zhang Z;
XX WPI: 2003-102117/09.
DR P-PSDB; ABU13920.
XX
PT Novel secreted and transmembrane polypeptide for modulating biological
PT activity of cell expressing the polypeptide, identifying agonists or
PT antagonists of polypeptide, and as molecular weight markers -
XX
PS Claim 2; Fig 150; 649pp; English.


```
PR 30-MAR-2000; 2000WO-US08439.
PR 15-MAY-2000; 2000WO-US13358.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 11-AUG-2000; 2000WO-US22031.
PR 23-AUG-2000; 2000WO-US23522.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 16-JUN-1997; 97US-043787P.
PR 17-OCT-1997; 97US-062250P.
PR 12-NOV-1997; 97US-065186P.
PR 13-NOV-1997; 97US-065311P.
PR 24-NOV-1997; 97US-066770P.
PR 25-FEB-1998; 98US-075945P.
PR 20-MAR-1998; 98US-078910P.
PR 28-APR-1998; 98US-083322P.
PR 07-MAY-1998; 98US-084600P.
PR 02-JUN-1998; 98US-087106P.
PR 02-JUN-1998; 98US-087607P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088021P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088026P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088030P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088738P.
PR 05-JUN-1998; 98US-088742P.
PR 10-JUN-1998; 98US-088810P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088858P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089440P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089532P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089599P.
PR 17-JUN-1998; 98US-089600P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089801P.
PR 18-JUN-1998; 98US-089907P.
PR 18-JUN-1998; 98US-089908P.
PR 28-AUG-2001; 2001US-0941992.
XX
PA (GETH ) GENENTECH INC.
XX
PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
PI Ferrara N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Kljavin IJ, Napier MA, Pan J, Paoni NF;
```

```
PI Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WT;
PI Zhang Z;
XX
DR WPI; 2003-066810/06.
DR P-PSDB; ABUI0875.
XX
PT Novel secreted and transmembrane polypeptide for modulating biological
PT activity of cell expressing the polypeptide, identifying agonists or
PT antagonists of polypeptide, and as molecular weight markers
XX
PS Claim 2; Fig 150; 655pp; English.
XX
CC The invention relates to a secreted and transmembrane polypeptide, termed
CC PRO polypeptide, and the polynucleotide encoding it. The polypeptide is
CC useful for detecting PRO polypeptides and for linking a bioactive
CC molecule to a cell expressing the above polypeptides, where the bioactive
CC molecule is a toxin, radiolabel or an antibody. The bioactive material
CC causes the death of the cell. The polypeptide is useful for identifying
CC agonists or antagonists of the PRO polypeptide, for preparing variants of
CC PRO, as a molecular weight marker for protein electrophoresis purposes
CC and the PRO polynucleotide is useful for recombinantly expressing those
CC markers. The polynucleotide is also useful as a hybridisation probe, in
CC chromosome and gene mapping, in generation of antisense RNA and DNA, in
CC the preparation of PRO polypeptide, for generating transgenic animals or
CC knockout animals which in turn are useful in the development and
CC screening of therapeutically useful reagents, to construct hybrid
CC probes for mapping the gene which encodes PRO and for the genetic
CC analysis of individuals with genetic disorders, in gene therapy, for
CC chromosome identification, as a chromosome marker and for generating
CC probes for PCR, Northern analysis, Southern analysis and Western
CC analysis. This sequence represents a human PRO polynucleotide of the
CC invention.
XX
SQ Sequence 992 BP; 235 A; 301 C; 282 G; 174 T; 0 other;

Query Match 100.0%; Score 992; DB 25; Length 992;
Best Local Similarity 100.0%; Pred. No. 2.7e-236;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCAAGGCGCGTCCCTGGGCTCTTCTGCTTGGCCGTGCTGGCTGCCAGCAGCTTCTCCAAAG 60
DB 1 GGCAAGGCGCGTCCCTGGGCTCTTCTGCTTGGCCGTGCTGGCTGCCAGCAGCTTCTCCAAAG 60
QY 61 GCATGGGGCTCCCTGGGCTCTTCTGCTTGGCCGTGCTGGCTGCCAGCAGCTTCTCCAAAG 120
DB 61 GCATGGGGCTCCCTGGGCTCTTCTGCTTGGCCGTGCTGGCTGCCAGCAGCTTCTCCAAAG 120
QY 121 CACGGAGGAGAAATTTACCCCTGTGGTCTCCATTGCTTCCATTGCTTCCAAAGTCTTGGAAAGTTTCC 180
DB 121 CACGGAGGAGAAATTTACCCCTGTGGTCTCCATTGCTTCCATTGCTTCCAAAGTCTTGGAAAGTTTCC 180
QY 181 CCAAGGCGCGTGGTGCTCATACCTGCTGTGCACCCAGCCAGCCACCCCATCACCT 240
DB 181 CCAAGGCGCGTGGTGCTCATACCTGCTGTGCACCCAGCCAGCCACCCCATCACCT 240
QY 241 ATTCCCTCTGTGAAACCAAGAACATCAAGTGTCGCAAGAGGTGGTGAACACCCAGC 300
DB 241 ATTCCCTCTGTGAAACCAAGAACATCAAGTGTCGCAAGAGGTGGTGAACACCCAGC 300
QY 301 CGGCGCTCTTCAACCTCAAGTCACTCAAGTCCAGTCCAGCTCCAGCTTCTTCTTCT 360
DB 301 CGGCGCTCTTCAACCTCAAGTCACTCAAGTCCAGTCCAGCTTCTTCTTCTTCTTCT 360
QY 361 GCGGGCGTCTCCACTCAGGTGCCATGTGCAGAGTGCCAGGCTACAGATGACTCTGGG 420
DB 361 GCGGGCGTCTCTCCACTCAGGTGCCATGTGCAGAGTGCCAGGCTACAGATGACTCTGGG 420
QY 421 AGCTGTGTCCTCAAGCCAGTCTTGTAGTGGGCCCACTTCACTCTGAGGACAGAGGGG 480
DB 421 AGCTGTGTCCTCAAGCCAGTCTTGTAGTGGGCCCACTTCACTCTGAGGACAGAGGGG 480
QY 481 CAGGCCCGAGGTGGAGATGATCTGCAGGCTCTTGGGCGAGCCACCTATCACAACA 540
DB 481 CAGGCCCGAGGTGGAGATGATCTGCAGGCTCTTGGGCGAGCCACCTATCACAACA 540
```

```
Db      481  CAGCCCCAGGTTGGAGATGATCTGCCAGGCGTCTCGGGCAGCCCACTATCACCAACA 540
QY      541  GCTGTATCGGAAGAGTGGCAGAGTCCACCTCGAGCAGAGACCATGCCACAGCGACGCTG 600
Db      541  GCCTGATCGGAAGAGTGGCAGAGTCCACCTCGAGCAGAGACCATGCCACAGCGACGCTG 600
QY      601  CCAACTTCCTCTCTCTCCGAGCAGCAGACATCGGACTGGTTCTGGTGCAGGCTGCAAAACA 660
Db      601  CCAACTTCCTCTCTCTCCGAGCAGCAGACATCGGACTGGTTCTGGTGCAGGCTGCAAAACA 660
QY      661  ACCGCAATGTCACAGCAGCGCCCTCACAGTGTGCCCGCAGGTGTGACCAAGATGG 720
Db      661  ACCGCAATGTCACAGCAGCGCCCTCACAGTGTGCCCGCAGGTGTGACCAAGATGG 720
QY      721  AGGACTGGCAGGTCCTCTCGAGAGGCCCACTCTTGCCTTGCCTTCCGCTCTACAGGAGCACCC 780
Db      721  AGGACTGGCAGGTCCTCTCGAGAGGCCCACTCTTGCCTTGCCTTCCGCTCTACAGGAGCACCC 780
QY      781  GCCGTCTGATGAAGAGAGTTTGGGGGTTTCAGGATAGGGAATGGGAGGTTCAGAGGAC 840
Db      781  GCCGTCTGATGAAGAGAGTTTGGGGGTTTCAGGATAGGGAATGGGAGGTTCAGAGGAC 840
QY      841  GCAAGCAGCAGCATGTAGAATGAACCGTCCAGAGAGCCCAAGCAGCAGGACTGCA 900
Db      841  GCAAGCAGCAGCATGTAGAATGAACCGTCCAGAGAGCCCAAGCAGCAGGACTGCA 900
QY      901  GGCCATCAGCGTCACTGTTCTGATTTGGAGTTTCATGCAAAATGAGTGTGTTTAGCTGC 960
Db      901  GGCCATCAGCGTCACTGTTCTGATTTGGAGTTTCATGCAAAATGAGTGTGTTTAGCTGC 960
QY      961  TCTTGCACAAAAAATAAAAAAAAAAAAAAAAAAAAAA 992
Db      961  TCTTGCACAAAAAATAAAAAAAAAAAAAAAAAAAAAA 992

RESULT 12
AAZ65030
ID  AAZ65030 standard; cDNA; 991 BP.
XX
AC  AAZ65030;
XX
DT  05-APR-2000 (first entry)
XX
DE  Membrane-bound protein PRO809 encoding cDNA.
XX
KW  Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;
pharmaceutical; receptor immunoadhesin; gene mapping; ss.
XX
OS  Homo sapiens.
XX
PN  WO963088-A2.
XX
PD  09-DEC-1999.
XX
PF  02-JUN-1999; 99WO-US12252.
XX
PR  02-JUN-1998; 98US-0087607.
PR  02-JUN-1998; 98US-0087609.
PR  02-JUN-1998; 98US-0087759.
PR  03-JUN-1998; 98US-0087827.
PR  04-JUN-1998; 98US-0088021.
PR  04-JUN-1998; 98US-0088025.
PR  04-JUN-1998; 98US-0088028.
PR  04-JUN-1998; 98US-0088029.
PR  04-JUN-1998; 98US-0088030.
PR  04-JUN-1998; 98US-0088033.
PR  04-JUN-1998; 98US-0088326.
PR  05-JUN-1998; 98US-0088167.
PR  05-JUN-1998; 98US-0088202.
PR  05-JUN-1998; 98US-0088212.
PR  05-JUN-1998; 98US-0088217.
PR  09-JUN-1998; 98US-0088655.
PR  10-JUN-1998; 98US-0088722.
PR  10-JUN-1998; 98US-0088730.
PR  10-JUN-1998; 98US-0088734.
PR  10-JUN-1998; 98US-0088738.
PR  10-JUN-1998; 98US-0088740.
PR  10-JUN-1998; 98US-0088741.
PR  10-JUN-1998; 98US-0088742.
PR  10-JUN-1998; 98US-0088810.
PR  10-JUN-1998; 98US-0088811.
PR  10-JUN-1998; 98US-0088824.
PR  10-JUN-1998; 98US-0088825.
PR  10-JUN-1998; 98US-0088826.
PR  11-JUN-1998; 98US-0088858.
PR  11-JUN-1998; 98US-0088861.
PR  11-JUN-1998; 98US-0088863.
PR  11-JUN-1998; 98US-0088876.
PR  12-JUN-1998; 98US-0089090.
PR  12-JUN-1998; 98US-0089105.
PR  16-JUN-1998; 98US-0089440.
PR  16-JUN-1998; 98US-0089512.
PR  16-JUN-1998; 98US-0089514.
PR  17-JUN-1998; 98US-0089532.
PR  17-JUN-1998; 98US-0089538.
PR  17-JUN-1998; 98US-0089598.
PR  17-JUN-1998; 98US-0089599.
PR  17-JUN-1998; 98US-0089600.
PR  17-JUN-1998; 98US-0089653.
PR  18-JUN-1998; 98US-0089801.
PR  18-JUN-1998; 98US-0089907.
PR  18-JUN-1998; 98US-0089908.
PR  19-JUN-1998; 98US-0089947.
PR  19-JUN-1998; 98US-0089948.
PR  19-JUN-1998; 98US-0089952.
PR  22-JUN-1998; 98US-0090246.
PR  22-JUN-1998; 98US-0090252.
PR  22-JUN-1998; 98US-0090254.
PR  23-JUN-1998; 98US-0090349.
PR  23-JUN-1998; 98US-0090355.
PR  24-JUN-1998; 98US-0090429.
PR  24-JUN-1998; 98US-0090431.
PR  24-JUN-1998; 98US-0090435.
PR  24-JUN-1998; 98US-0090444.
PR  24-JUN-1998; 98US-0090445.
PR  24-JUN-1998; 98US-0090461.
PR  24-JUN-1998; 98US-0090472.
PR  24-JUN-1998; 98US-0090535.
PR  24-JUN-1998; 98US-0090538.
PR  24-JUN-1998; 98US-0090540.
PR  24-JUN-1998; 98US-0090557.
PR  25-JUN-1998; 98US-0090676.
PR  25-JUN-1998; 98US-0090678.
PR  25-JUN-1998; 98US-0090688.
PR  25-JUN-1998; 98US-0090690.
PR  25-JUN-1998; 98US-0090691.
PR  25-JUN-1998; 98US-0090694.
PR  25-JUN-1998; 98US-0090695.
PR  26-JUN-1998; 98US-0090696.
PR  26-JUN-1998; 98US-0090862.
PR  26-JUN-1998; 98US-0090863.
PR  01-JUL-1998; 98US-0091358.
PR  01-JUL-1998; 98US-0091360.
PR  02-JUL-1998; 98US-0091478.
PR  02-JUL-1998; 98US-0091486.
PR  02-JUL-1998; 98US-0091519.
PR  02-JUL-1998; 98US-0091626.
PR  02-JUL-1998; 98US-0091628.
PR  02-JUL-1998; 98US-0091633.
PR  02-JUL-1998; 98US-0091646.
PR  02-JUL-1998; 98US-0091673.
PR  07-JUL-1998; 98US-0091978.
PR  07-JUL-1998; 98US-0091982.
PR  09-JUL-1998; 98US-0092182.
PR  10-JUL-1998; 98US-0092472.
```


20-JUL-1998; 98US-0093339.
30-JUL-1998; 98US-0094651.
04-AUG-1998; 98US-0095282.
04-AUG-1998; 98US-0095282.
04-AUG-1998; 98US-0095285.
04-AUG-1998; 98US-0095301.
04-AUG-1998; 98US-0095302.
04-AUG-1998; 98US-0095318.
04-AUG-1998; 98US-0095321.
04-AUG-1998; 98US-0095325.
10-AUG-1998; 98US-0095316.
10-AUG-1998; 98US-0095929.
10-AUG-1998; 98US-0096012.
11-AUG-1998; 98US-0096143.
11-AUG-1998; 98US-0096146.
12-AUG-1998; 98US-0096329.
17-AUG-1998; 98US-0096757.
17-AUG-1998; 98US-0096766.
17-AUG-1998; 98US-0096768.
17-AUG-1998; 98US-0096773.
17-AUG-1998; 98US-0096791.
17-AUG-1998; 98US-0096867.
17-AUG-1998; 98US-0096891.
17-AUG-1998; 98US-0096894.
17-AUG-1998; 98US-0096895.
17-AUG-1998; 98US-0096897.
18-AUG-1998; 98US-0096949.
18-AUG-1998; 98US-0096950.
18-AUG-1998; 98US-0096959.
18-AUG-1998; 98US-0096960.
18-AUG-1998; 98US-0096966.
18-AUG-1998; 98US-0097022.
19-AUG-1998; 98US-0097141.
20-AUG-1998; 98US-0097218.
20-AUG-1998; 98US-0097661.
26-AUG-1998; 98US-0097951.
26-AUG-1998; 98US-0097952.
26-AUG-1998; 98US-0097954.
26-AUG-1998; 98US-0097955.
26-AUG-1998; 98US-0097971.
26-AUG-1998; 98US-0097974.
26-AUG-1998; 98US-0097978.
26-AUG-1998; 98US-0097979.
26-AUG-1998; 98US-0097986.
26-AUG-1998; 98US-0098014.
31-AUG-1998; 98US-0098525.
16-SEP-1998; 98US-0100634.
12-JAN-1999; 99US-0115565.
(GETH) GENENTECH INC.
Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK;
Wood WL, Yuan J;
WPI; 2000-072883/06.
P-PSDB; AAY66691.
Membrane-bound proteins and related nucleotide sequences -
Claim 2; Fig 150; 822pp; English.
The invention provides membrane-bound PRO polypeptides and
polynucleotides encoding them. The PRO sequences of the invention were
identified based on extracellular domain homology screening. The PRO
sequences have homology with proteins including LDL receptors, TIE
ligands and various enzymes. The membrane-bound proteins and receptor
molecules are useful as pharmaceutical and diagnostic agents. Receptor
immunoadhesins, for instance, can be used as therapeutic agents to block
receptor-ligand interactions. The membrane-bound proteins can also be
employed for screening of potential peptide or small molecule inhibitors
of the relevant receptor/ligand interaction. The PRO encoding sequences
are useful as hybridization probes, in chromosome and gene mapping and in
the generation of antisense RNA and DNA. PRO nucleic acid sequences
will also be useful for the preparation of PRO polypeptides, especially
by recombinant techniques.

QY	157	CTTACAAAGTCTCTGGAAGTTTTCCCCAAAGCCGCTGGGTGCTCATAAACCTGCTGTGCAC	211
DB	121	CTTACAAAGTCTCTGGAAGTTTTCCCCAAAGCCGCTGGGTGCTCATAAACCTGCTGTGCAC	180
QY	217	CCCAGCCACCCGCGCCATCACCTATTCCCTCTGTGTGGAACCAAGAACAATCAAGGTGGCCA	276
DB	181	CCCAGCCACCCGCGCCATCACCTATTCCCTCTGTGTGGAACCAAGAACAATCAAGGTGGCCA	240
QY	277	AGAGGTGGTGAAGACCCACAGACCGGGCTCTTCAACCTCAACGTCACTCAAGTCCA	336
DB	241	AGAGGTGGTGAAGACCCACAGACCGGGCTCTTCAACCTCAACGTCACTCAAGTCCA	300
QY	337	GTCCAGACCTGCTCACCTACTTCTGCGGGGGTCTCCACCTCAGGTGCCATGTGGACA	396
DB	301	GTCCAGACCTGCTCACCTACTTCTGCGGGGGTCTCCACCTCAGGTGCCATGTGGACA	360
QY	397	GTGCCAGGCTACAGATGCATCTGGGAGCTGTGGTCAAGCCAGTGTCTGAGTTCGGGGCCA	456
DB	361	GTGCCAGGCTACAGATGCATCTGGGAGCTGTGGTCAAGCCAGTGTCTGAGTTCGGGGCCA	420
QY	457	ACTTCACTCTGACGACAGAGGGCGGCCCCAGGTGGAGATGATCTGCCAGGCGTCT	516
DB	421	ACTTCACTCTGACGACAGAGGGCGGCCCCAGGTGGAGATGATCTGCCAGGCGTCT	480
QY	517	CGGGCAGCCCCACTTATCACCAACAGCCTGATCGGGAAGGATGGCAGGTCCACCTGCAGC	576
DB	481	CGGGCAGCCCCACTTATCACCAACAGCCTGATCGGGAAGGATGGCAGGTCCACCTGCAGC	540
QY	577	AGAGCAATGSCCA CAGGAGCGCTGCCAACTTCTCTTCTGCCAGCCAGACATCGGACT	636
DB	541	AGAGCAATGSCCA CAGGAGCGCTGCCAACTTCTCTTCTGCCAGCCAGACATCGGACT	600
QY	637	GGTCTCTGTCAGGCTGCAGCAACCAAGCCATGTCCAGACAGCGCCCTCACTAGTGTGC	696
DB	601	GGTCTCTGTCAGGCTGCAGCAACCAAGCCATGTCCAGACAGCGCCCTCACTAGTGTGC	660
QY	697	CC-----	698
DB	661	CCCCAGGAGGGTTGCCCAGGGCACCACCATCTGTGCTGGTGGCAGCTTGCTTCCACTG	720
QY	699	-----CCAGGTGGTACCAGGAAGATG	719
DB	721	CGGCCATCATCTCCAGGATGCTGGGCTGGACCACGTGGGCCAGGTGGTGACCAGAGATG	780
QY	720	GAGGACTGGCAGGGTCCCTTGGAGAGCCCCATCTCTTGCCTTGGCGTCTACAGAGACAC	779
DB	781	GAGGACTGGCAGGGTCCCTTGGAGAGCCCATCTCTTGCCTTGGCGTCTACAGAGACAC	840
QY	780	CGCGCTCTGAGTGAAGAGGAGTTTGGGGGGTTTCAAGGATAGGGAATGGGAGGTCAAGAGGA	839
DB	841	CGCGCTCTGAGTGAAGAGGAGTTTGGGGGGTTTCAAGGATAGGGAATGGGAGGTCAAGAGGA	900
QY	840	CGCAAGCAGCAGCGCATGTAGATGAACCTTCAGAGAGCCAGACACGGCAGAGGACTGC	899
DB	901	CGCAAGCAGCAGCGCATGTAGATGAACCTTCAGAGAGCCAGACACGGCAGAGGACTGC	960
QY	900	AGSCCATCAGCGTGCACTGTTCTGATTTGGAGTTTCATGCAAAATGATGTGTTTTAGCTG	959
DB	961	AGSCCATCAGCGTGCACTGTTCTGATTTGGAGTTTCATGCAAAATGATGTGTTTTAGCTG	1020
QY	960	CTCTTGCCACAAAAAATAAAAAAAAAA	986
DB	1021	CTCTTGCCACAAAAAATAAAAAAAAAA	1047
RESULT 14			
AAH99131			
ID AAH99131 standard; cDNA; 935 BP.			
XX AAH99131;			
XX			
12-OCT-2001 (first entry)			
XX			

DE Human EST-derived coding sequence SEQ ID NO: 988.
 XX Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
 KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
 KW diagnostics; forensic test; gene mapping; genetic disorder;
 KW biodiversity; gene therapy; nutrition; ss.
 XX
 OS Homo sapiens.
 XX WO200154477-A2.
 PN
 XX
 XX 02-AUG-2001.
 PD
 PF 25-JAN-2001; 2001WO-US02687.
 XX
 PR 25-JAN-2000; 2000US-0491404.
 PR 17-JUL-2000; 2000US-0617746.
 PR 03-AUG-2000; 2000US-0631451.
 PR 15-SEP-2000; 2000US-0663870.
 XX
 XX (HYSE-) HYSEQ INC.
 PA
 XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
 PI Cao Y, Drmanac RA, Zhang J, Werhman T;
 XX
 DR WPI; 2001-476164/51.
 DR P-PSDB; AM24472.
 XX
 PT Isolated polypeptide for treatment of diseases, diagnostics, raising
 PT antibodies and research use -
 XX
 PS Claim 1; Page 770; 1275pp; English.
 CC The present invention provides the protein and coding sequences of novel
 CC proteins from a variety of organisms, including human, dog, cat, horse,
 CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
 CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
 CC from the organism of interest. They can be used in diagnostics,
 CC forensics, gene mapping, identification of mutations, to assess
 CC biodiversity and for nutritional purposes. The present sequence is a cDNA
 CC of the invention.
 XX
 SQ Sequence 935 BP; 206 A; 286 C; 268 G; 175 T; 0 other;
 Query Match 79.7%; Score 790.2; DB 22; Length 935;
 Best Local Similarity 94.6%; Pred. No. 3.2e-186;
 Matches 920; Conservative 0; Mismatches 8; Indels 45; Gaps 8;
 QY 9 CCAGAACTAGGAGGTTCTACTGCCGAGCAGAGGCCCTACACCCACCGAGGATGGG 68
 DB 1 CCAGGAACTAGGAGGTTCTACTGCCGAGCAGAGGCCCTACACCCACCGAGGATGGG 60
 QY 69 CTCCTGGGCTGTTCTGTTGGCGTGTGCTGCCAGCAGCTTCTCCAGGACGGAG 128
 DB 61 CTCCTGGGCTGTTCTGTTGGCGTGTGCTGCCAGCAGCTTCTCCAGGACGGAG 120
 QY 129 GAAGAAATACCCCTGTGGTCTCCATTGCCCTACAAAGTCTCGGAAGTTTCCCAAGGC 188
 DB 121 GAAGAAATACCCCTGTGGTCTCCATTGCCCTACAAAGTCTCGGAAGTTTCCCAAGGC 180
 QY 189 CGCTGGGTGTCATAAAGTGTGTGCAACCCAGGCCACCGCCCATCACTATTTCCTC 248
 DB 181 CGCTGGGTGTCATAAAGTGTGTGCAACCCAGGCCACCGCCCATCACTATTTCCTC 240
 QY 249 TGTGGACCAAGACATCAAGTGTGGCCAGAGAGTGTGAAGACCCAGGACGGCTCC 308
 DB 241 TGTGGACCAAGACATCAAGTGTGGCCAGAGAGTGTGAAGACCCAGGACGGCTCC 300
 QY 309 TTCAACCTCAACGTCACACTCAAGTGTGCAACCCAGGCCACCGCCCATCACTATTTCCTC 368
 DB 301 TTCAACCTCAACGTCACACTCAAGTGTGCAACCCAGGCCACCGCCCATCACTATTTCCTC 360
 QY 369 TCCTCCACCTCAGGTGGCCATGTGGACAGTGTGCCAGGTGCCAGGTACAGATGCACCTGGAGCTGTGG 428

DB 361 TCCTCCACCTCAGGTGGCCATGTGGACAGTCCAGGCTACAGATGCATCGGAGCTGTGG 420
 QY 429 TCCAAGCCAGTGTCTGAGTCTGGGGCCAACTTCACTCTGACGAGCAGAGGGGCGAGCCCC 488
 DB 421 TCCA-----GACAGAGGGGCGAGCCCC 442
 QY 489 AGGTGAGATGATCTCCAGGCGCTCTGGGCGAGCCCACTATCAACAACAGCTGATC 548
 DB 443 AGGTGAGATGATCTCCAGGCGCTCTGGGCGAGCCCACTATCAACAACAGCTGATC 502
 QY 549 GGGAGGATGGGAGGTCCACCTTCAGCAGAGACCATGCCACGCGAGCTGCCAAC-TT 607
 DB 503 GGGAGGATGGGAGGTCCACCTTCAGCAGAGACCATGCCACGAGGCGCTGCCAAC-TT 562
 QY 608 CTCCTTCTCCGAGCCAGACATCGGAC-TGGTTTCTGGTCCAGGC-TGCAAAACAAGGCC 665
 DB 563 CTCCTTCTCCGAGCCAGACATCGGACTTGGTTCTGGTCCAGGCTTGCAAAACAAGGCC 622
 QY 666 AATG-TCCAGCAGAGCGCCCTCAAGTGTGGTGGCCCGCAGGTGGTGA-CCAGAAATGGAGG 723
 DB 623 AATGTTCCAGCAGAGCGCCCTCAAGTGTGGTGGCCCGCAGGTGGTGA-CCAGAAATGGAGG 682
 QY 724 ACTGGCAGGCT-CCCTGGAGAGCCCATCTTGGCTTGGCGCTCTACAGGACACCGC 782
 DB 683 ACTGGCAGGCTCCCTGGAGAGCCCATCTTGGCTTGGCGCTCTACAGGACACCGC 742
 QY 783 CGTCTGAGTCAAGAGGAG-TTTGGGGGGTTTCAGGATAGGGAATGGGAGGTTCAGAGGAGC 841
 DB 743 CGTCTGAGTCAAGAGGAGTTTGGGGGGTTTCAGGATAGGGAATGGGAGGTTCAGAGGAGC 802
 QY 842 CAAGCAGCAGCCATGTAGATGACCGTCCAGAGAGCCAGCAGCAGGACGCTGCAG 901
 DB 803 CAAGCAGCAGCCATGTAGATGACCGTCCAGAGAGCCAGCAGCAGGACGCTGCAG 862
 QY 902 GCCATCAGCGTGCACCTTCTGATTTGGAGTTTCATGCAAAATGAGTGTGTTTAGCTGCT 961
 DB 863 GCCATCAGCGTGCACCTTCTGATTTGGGGTTTCATGCAAAATGAGTGTGTTTAGCTGCT 922
 QY 962 CTTGCCCAAAAA 974
 DB 923 CTTGCCCAAAAA 935
 RESULT 15
 AAC74260
 ID AAC74260 standard; cDNA; 1047 BP.
 XX
 AC AAC74260;
 XX
 DT 02-FEB-2001 (first entry)
 XX
 DE Human secreted protein gene 38 SEQ ID NO:48.
 XX
 KW Human; secreted protein; immunosuppressive; antiarthritic; antirheumatic;
 KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
 KW neotropic; neuroprotective; antibacterial; virucide; fungicide; neoplasm;
 KW ophthalmological; autoimmune disease; rheumatoid arthritis; angiogenesis;
 KW hyperproliferative disorder; cardiovascular disorder; infection;
 KW cerebrovascular disorder; nervous system disorder; ocular disorder;
 KW wound healing; chemotaxis; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200056754-A1.
 XX
 PD 28-SEP-2000.
 XX
 PF 16-MAR-2000; 2000WO-US06792.
 XX
 PR 19-MAR-1999; 99US-0125362.
 PR 10-DEC-1999; 99US-0169980.
 XX

CC therapy. The genes are isolated from a range of human tissues disclosed
CC in the specification. The nucleic acids, proteins, antibodies and
CC (ant)agonists are useful in the diagnosis, treatment and prevention of:
CC (a) cancer, e.g. breast and ovarian cancer and other cancers of the
CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
CC lung or urogenital; (b) immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; (c) cardiovascular diseases such as
CC myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g.
CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
CC bacterial, fungal and parasitic infections.

XX SQ Sequence 1047 BP; 238 A; 323 C; 299 G; 184 T; 3 other;

Query Match 78.0%; Score 773.6; DB 25; Length 1047;
Best Local Similarity 87.1%; Pred. No. 4.4e-182;
Matches 941; Conservative 1; Mismatches 0; Indels 138; Gaps 3;

QY 10 CAGGAACCTAGAGGTTCTCTACCTGCCGAGCAGA-GGCCCTTACACCCACCGAGGATGGG 68
DB 1 CAGGAACCTAGAGGTTCTCTACCTGCCGAGCAGA-GGCCCTTACACCCACCGAGGATGGG 60
QY 69 CTCCTGGGCTGTTCTGCTGGCGTGTCTGCTGCCAGAGCTTCTCCAGGACAGGGAG 128
DB 61 CTCCTGGGCTGTTCTGCTGGCGTGTCTGCTGCCAGAGCTTCTCCAGGACAGGGAG 120
QY 129 GAAGAAATTAACCCCTGTGGTCTCATGTGCTTCAAGTCTCTGGAAGTTTTCGCCAAGGC 188
DB 121 GAAGAAATTAACCCCTGTGGTCTCATGTGCTTCAAGTCTCTGGAAGTTTTCGCCAAGGC 180
QY 189 CGCTGGGTGCTATAACCTGCTGTGACCCGACCCACCGCCCATCATCTATTCCCTC 248
DB 181 CGCTGGGTGCTATAACCTGCTGTGACCCGACCCACCGCCCATCATCTATTCCCTC 240
QY 249 TGTGGACCAAGACATCAGGTGGCCAGAAAGTGTGAAGACCCACGAGCGGCTCC 308
DB 241 TGTGGACCAAGACATCAGGTGGCCAGAAAGTGTGAAGACCCACGAGCGGCTCC 300
QY 309 TTCAACCTCAACCTCACTCAAGTCCAGTCCAGACCTGCTCACTTCTGCCGGCG 368
DB 301 TTCAACCTCAACCTCACTCAAGTCCAGTCCAGACCTGCTCACTTCTGCCGGCG 360
QY 369 TCTTCACTCAGGTGCCATGTGACAGTGCAGGTCAGGTCACAGTCACTGGGAGCTGTGG 428
DB 361 TCTTCACTCAGGTGCCATGTGACAGTGCAGGTCAGGTCACAGTCACTGGGAGCTGTGG 420
QY 429 TCCAAGCCAGTGTGAGCTGCGGGCCAACTTCACTCTGACGACAGGGGGCAGGCCCC 488
DB 421 TCCA-----GACAGGGGGCAGGCCCC 442
QY 489 AGGTGGAGATGATGTGCCAGGCGTCTCGGGCAGGCCCACTATCAACACGCTTGATC 548
DB 443 AGGTGGAGATGATGTGCCAGGCGTCTCGGGCAGGCCCACTATCAACACGCTTGATC 502
QY 549 GGAAGAGTGGGAGGTTCACCTGACAGACACCATGACAGGACGCTGCGCACTTC 608
DB 503 GGAAGAGTGGGAGGTTCACCTGACAGACACCATGACAGGACGCTGCGCACTTC 562
QY 609 TCTTCTGCGAGCCAGACATCGGACTGTTCTGTTGTCAGGCTGCAACACGCGCAAT 668
DB 563 TCTTCTGCGAGCCAGACATCGGACTGTTCTGTTGTCAGGCTGCAACACGCGCAAT 622
QY 669 GTCCAGCACAGCGCCCTCACTGAGTGTGTC----- 697
DB 623 GTCCAGCACAGCGCCCTCACTGAGTGTGTC----- 682
QY 698 ----- 697
DB 683 GTGCTGTTGGACGCTTGGCTTCACTGCGGCCATCACCTCCAGGATGTGGGCTGGAC 742
QY 698 -----CCAGGTGGTACACAGATGGAGACTGGCAGGTCCCTCTGGAGACCCC 749

DB 743 CACGTGGGCCAGGTGGTGTGACCAAGATGGAGGACTGGCAGGTCCTCTGGAGAGCCCC 802
QY 750 ATCCTTGCCTTCCGCTCTACAGGACACCCCGCTGTGAGTGAAGAGGAGTTTGGGGG 809
DB 803 ATCCTTGCCTTCCGCTCTACAGGACACCCCGCTGTGAGTGAAGAGGAGTTTGGGGG 862
QY 810 TTCAGGATAGGGAATGGGAGGTTCAGAGGACGCAAGACGACGACCATGTAGAATGAACG 869
DB 863 TTCAGGATAGGGAATGGGAGGTTCAGAGGACGCAAGACGACGACCATGTAGAATGAACG 922
QY 870 TCAGAGAGCCAGACGCGGAGGACTGAGGCGCATCAGCGTGCACCTGTTCTGTTATTTGG 929
DB 923 TCAGAGAGCCAGACGCGGAGGACTGAGGCGCATCAGCGTGCACCTGTTCTGTTATTTGG 982
QY 930 AGTTTCATGCAAAATGAGTGTGTTTGTAGTCTCTTCCACAAAAAAGGAGGAGGAGG 989
DB 983 AGTTTCATGCAAAATGAGTGTGTTTGTAGTCTCTTCCACAAAAAAGGAGGAGGAGG 1042

RESULT 18

ABZ36822

ID ABZ36822 standard; cDNA; 502 BP.

XX AC ABZ36822;

XX XX 21-FEB-2003 (first entry)

XX XX Human GENSET coding sequence SEQ ID 695.

XX Cytostatic; antiinflammatory; nootropic; neuroprotective; cardiac;
XX gastrointestinal; gene therapy; GENSET; heavy metal toxicity; cancer;
XX inflammatory disease; immune disorder; neuromuscular; toxicity;
XX central nervous system; cardiovascular; gastrointestinal; gene; ss.

OS Homo sapiens.

XX WO200283898-A1.

XX PD 24-OCT-2002.

XX PF 18-APR-2001; 2001WO-IB00914.

XX PR 18-APR-2001; 2001WO-IB00914.

XX PA (GENSET) GENSET.

XX Bejanin S, Tanaka H, Dumas Milne Edwards J, Jobert S, Giordano J;
XX WPI; 2003-075548/07.

XX New GENSET polynucleotides and polypeptides, useful for treating heavy
XX metal toxicity, cancer, inflammatory diseases, immune disorders, and
XX the neuromuscular, CNS, cardiovascular or gastrointestinal effects of
XX the toxicity -

XX Claim 12; Page 656; 735pp; English.

XX The present invention relates to novel GENSET polynucleotides
XX (ABZ36404-ABZ36911) encoding polypeptides (ABP75963-ABP76368). The
XX polynucleotides and polypeptides are useful in screening and diagnostic
XX assays for abnormal GENSET expression and/or biological activity. They
XX are also useful for screening of compounds for treating or preventing
XX GENSET-related disorders, such as heavy metal toxicity, cancer,
XX inflammatory diseases, immune disorders, and the neuromuscular, central
XX nervous system (CNS), cardiovascular or gastrointestinal effects of the
XX toxicity. The polynucleotides are useful for constructing or expanding
XX chromosome maps.

XX SQ Sequence 502 BP; 105 A; 169 C; 135 G; 92 T; 1 other;

Query Match 50.6%; Score 501.6; DB 25; Length 502;
Best Local Similarity 99.8%; Pred. No. 1.1e-114;
Matches 501; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY 40 AGAGGCCCTACACCCAGGAGGATGGGGTCTCCCTGGGTGTTCTGCTTGGCGGTCTGG 99
Db 1 AGAGGCCCTACACCCAGGAGGATGGGGTCTCCCTGGGTGTTCTGCTTGGCGGTCTGG 60
QY 100 CTGCCAGAGCTTCTCCAGGACGAGGAGAGAAATTACCCCTGGTCTCCATTGCTT 159
Db 61 CTGCCAGAGCTTCTCCAGGACGAGGAGAGAAATTACCCCTGGTCTCCATTGCTT 120
QY 160 ACAAGTCTGGAAGTTTCCCAAGGCGCTGGTCTATAACCTGCTGTGCACCCC 219
Db 121 ACAAGTCTGGAAGTTTCCCAAGGCGCTGGTCTATAACCTGCTGTGCACCCC 180
QY 220 AGCCACACCGCCCATCACTATTCCTCTGTGGAACCAAGAACATCAAGGTGGCAAGA 279
Db 181 AGCCACACCGCCCATCACTATTCCTCTGTGGAACCAAGAACATCAAGGTGGCAAGA 240
QY 280 AGGTGTGAAGACCCAGGAGCGGCTCTTCAACCTCAACGTCACACTCAAGTCCAGTC 339
Db 241 AGGTGTGAAGACCCAGGAGCGGCTCTTCAACCTCAACGTCACACTCAAGTCCAGTC 300
QY 340 CAGACTGCTACCTACTTCTCCGGGGTCTCCACCTCAGGTGCCATGTGGACAGTG 399
Db 301 CAGACTGCTACCTACTTCTCCGGGGTCTCCACCTCAGGTGCCATGTGGACAGTG 360
QY 400 CCAGGCTACAGTGCACCTGGGAGCTGTGTCCAAGCAGTGTCTGAGTGGGGCAACT 459
Db 361 CCAGGCTACAGTGCACCTGGGAGCTGTGTCCAAGCAGTGTCTGAGTGGGGCAACT 420
QY 460 TCACCTCTGAGACAGAGGGGAGGCGGCTCTTCAACCTCAACGTCAGATGATCTGCCAGGCTCCTCGG 519
Db 421 TCACCTCTGAGACAGAGGGGAGGCGGCTCTTCAACCTCAACGTCAGATGATCTGCCAGGCTCCTCGG 480
QY 520 GCAGGCCACCTATCAACCAAG 541
Db 481 GCAGGCCACCTATCAACCAAG 502

RESULT 19
AAS34076
ID AAS34076 standard; cDNA; 724 BP.
AC AAS34076;
DT 17-DEC-2001 (first entry)
DE Human cDNA encoding a novel foetal antigen, SEQ ID No 600.
XX
KW Human; foetal tissue antigen; ss; antiinflammatory; neuroprotective;
KW immunomodulator; cardiovascular; cytostatic; nephrothropic;
KW cardiovascular; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorder; breast neoplasm; cancer;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; food additive.
XX
OS Homo sapiens.
XX
FN WO200155312-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01321.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0225759.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
```

PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 01-DEC-2000; 2000US-0250160.
 PR 05-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 08-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 (HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-488782/53.
 P-PSDB; AAU21256.

New polynucleotides and polypeptides for diagnosing, treating, preventing or prognosing e.g. diseases or disorders of the nervous, musculoskeletal, excretory, gastrointestinal, reproductive, and respiratory systems -

Claim 1; SEQ ID No 600; 642pp; English.

The invention relates to novel nucleic acids encoding novel human foetal antigens. The nucleic acids and proteins are used to prevent, treat (e.g. by gene therapy) or ameliorate a medical condition in e.g. humans, mice,

CC rabbits, goats, horses, cats, dogs, chickens or sheep. They
 CC are also used in diagnosing a pathological condition or susceptibility
 CC to a pathological condition. The antibodies to the antigens can also
 CC be used in alleviating symptoms associated with the disorders and in
 CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
 CC immunosorbent assays (ELISA). Disorders which are diagnosed or treated
 CC include autoimmune diseases e.g. rheumatoid arthritis,
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
 CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi
 CC and ocular disorders e.g. corneal infection. The polypeptides can also
 CC be used to aid wound healing and epithelial cell proliferation, to
 CC prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, for supporting cell culture of primary tissues, to
 CC regenerate tissues and in chemotaxis. The polypeptides can also be used
 CC as a food additive or preservative to increase or decrease storage
 CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
 CC minerals, cofactors and other nutritional components. Numerous
 CC examples of diseases and disorders treated by the nucleic acids and
 CC proteins are given in the specification. The present sequence

Query Match 50.6%; Score 501.6; DB 22; Length 724;
 Best Local Similarity 98.0%; Pred. No. 1.2e-114;

Matches 545; Conservative 4; Mismatches 3; Indels 4; Gaps 4;

QY	131	AGAAATTACCCCTGCTGCTCCATTGCTTACAAAGTCTTGGAAAGTTTCCCAAGGCGG	190
DB	161	AGAAATTACCCCTGCTGCTCCATTGCTTACAAAGTCTTGGAAAGTTTCCCAAGGCGG	220
QY	191	CTGGGTGCTCATAAOCCTGCTGTGACCCAGCCACCCAGCCCATCACCTATTCCTCTG	250
DB	221	CTGGGTGCTCATAAOCCTGCTGTGACCCAGCCACCCAGCCCATCACCTATTCCTCTG	280
QY	251	TGGAACCAAGACATCAAGTGGCGCAAGAGTGGTGAAGACCCAGCGCGCTCCTT	310
DB	281	TGGAACCAAGACATCAAGTGGCGCAAGAGTGGTGAAGACCCAGCGCGCTCCTT	340
QY	311	CAACCTCAACGTCACACTCAAGTCCAGTCCAGACCTGCTCAGCTACTTCTGCGGCGTC	370
DB	341	CAACCTCAACGTCACACTCAAGTCCAGTCCAGACCTGCTCAGCTACTTCTGCGGCGTC	400
QY	371	CTCCACCTCAGTGGCCCATGTGGACAGTGCAGGGCTACAGATGCATGGGAGCTGTGTC	430
DB	401	CTCCACCTCAGTGGCCCATGTGGACAGTGCAGGGCTACAGATGCATGGGAGCTGTGTC	460
QY	431	CAAGCCAGTGTGAGCTGCGGGCCAACTTCACTCTGCAGGACAGAGGGGAGGCCAG	490
DB	461	CAAGCCAGTGTGAGCTGCGGGCCAACTTCACTCTGCAGGACAGAGGGGAGGCCAG	519
QY	491	GGTGAGATGATCTGCCAGGCGTCTCGGGGAGCCCACTATCACCAACAGCGCTGATCGG	550
DB	520	GGTGAGATGATCTGCCAGGCGTCTCGGGGAGCCCACTATCACCAACAGCGCTGATCGG	578
QY	551	GAAGGATGGGAGGTCACCTGTCAGCAGAGACCATGCCACAGGAGCTGCCAACTTCTC	610
DB	579	GAAGGATGGGAGGTCACCTGTCAGCAGAGACCATGCCACAGGAGCTGCCAACTTCTC	637
QY	611	CTTCTGCGGAGCCAGACATCGGACTGTTTC-TGGTGCAGGCTGCMAACACGCAATG	669
DB	638	CTTCTGCGGAGCCAGACATCGGACTGTTTC-TGGTGCAGGCTGCMAACACGCAATG	697
QY	670	TCCAGCACAGCGCCCT 685	
DB	698	TCCAGCACAGCGCCCT 713	

RESULT 20
 ABA20618

ID ABA20618 standard; DNA; 32220 BP.

XX
 AC ABA20618;
 XX

DT XX 23-JAN-2002 (first entry)
DE XX Human nervous system related polynucleotide SEQ ID NO 12949.
KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antiskilking; antianemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antifungal; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX OS Homo sapiens.
XX XX WO200159063-A2.
XX PD 16-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US01334.
XX PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0215647.
PR 07-JUL-2000; 2000US-0216680.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 05-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234224.
PR 25-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 26-SEP-2000; 2000US-0234998.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0242221.
PR 08-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.


```
Db 30 AAATGAGTGTGTTTGTAGCTGCTCTTGCCAC 1
RESULT 23
AAAX12009/c
ID AAX12009 standard; DNA; 150 BP.
AC AAX12009;
XX
DT 30-MAR-1999 (first entry)
XX
DE Human biallelic polymorphic DNA fragment EST376246a.
XX
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medication;
KW treatment; marker; ss.
XX
OS Homo sapiens.
XX
FN WO9820165-A2.
XX
PD 14-MAY-1998.
XX
PF 05-NOV-1997; 97WO-US20313.
XX
PR 06-NOV-1996; 96US-0030455.
XX
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
XX
PI Hudson T, Lander ES, Wang D;
XX
DR WPI; 1998-286974/25.
XX
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
XX
PS Claim 1; Page 212; 310pp; English.
XX
CC AAX10269-X12937 are human DNA fragments which contain biallelic
CC polymorphic markers which have been isolated using the primers
CC represented in AAX09121-X10268. The base occupying the polymorphic site
CC is indicated by the appropriate IUPAC-IVB ambiguity code. These fragments
CC can be used in methods for determining polymorphic forms in an individual
CC for use in e.g. forensics, paternity testing or for phenotypic typing for
CC diseases such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan
CC syndrome, muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease,
CC familial hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
XX
SQ Sequence 150 BP; 30 A; 46 C; 33 G; 40 T; 1 other;
Query Match 15.1%; Score 149.6; DB 19; Length 150;
Best Local Similarity 99.3%; Pred. No. 1.6e-27;
Matches 149; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 820 GGATGGGAGGTGAGAGGACGCGAAGCAGCAGCCATGTAGATGAACCGTCCAGAGC 879
Db 150 GGAATGGGAGGTGAGAGGACGCGAAGCAGCAGCCATGTAGATGAACCGTCCAGAGC 91
Qy 880 CAAGCAGGACGAGACTGCGAGGCCATCAGCGTCACTGTTGATTTGAGTTTCATGCA 939
Db 90 CAAGCAGGACGAGACTGCGAGGCCATCAGCGTCACTGTTGATTTGAGTTTCATGCA 31
```

PR 16-AUG-2001; 2001US-312903P.
 PR 10-SEP-2001; 2001US-318462P.
 PR 12-SEP-2001; 2001US-318770P.
 PR 27-SEP-2001; 2001US-325430P.
 PR 27-SEP-2001; 2001US-325681P.
 PR 18-OCT-2001; 2001US-330380P.
 PR 31-OCT-2001; 2001US-335301P.
 PR 14-NOV-2001; 2001US-332172P.
 PR 14-NOV-2001; 2001US-332271P.
 PR 14-NOV-2001; 2001US-332272P.
 PR 14-NOV-2001; 2001US-333184P.
 PR 14-NOV-2001; 2001US-333272P.
 PR 21-NOV-2001; 2001US-332094P.
 PR 03-DEC-2001; 2001US-337426P.
 PR 03-DEC-2001; 2001US-338092P.
 PR 04-DEC-2001; 2001US-337185P.
 PR 03-JAN-2002; 2002US-345705P.
 PR 07-MAR-2002; 2002US-0092900.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Padigar M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
 PI Zerhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
 PI Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;
 PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y;
 PI Anderson D, Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H;
 PI Alsobrook JP, Lepley DM, Rieger DK;
 XX
 DR WPI; 2002-723332/78.
 DR P-PSDB; ABU65174.
 XX
 PT NOVX polypeptides and polynucleotides, useful for preventing or
 PT treating a disorder associated with aberrant NOVX expression or
 PT activity e.g., cancer, hypertension, atherosclerosis, cardiomyopathy or
 PT bronchial asthma
 XX
 PS Claim 13; Page 368-369; 1103pp; English.
 XX
 CC This invention describes novel human NOVX polypeptides which have
 CC cytotstatic, cardiant, antiarteriosclerotic, antiasthmatic and
 CC hypotensive activity. Pharmaceutical compositions comprising the NOVX
 CC proteins or nucleic acid molecules or NOVX antibodies are useful for
 CC preventing or treating a disorder associated with aberrant NOVX
 CC expression or activity e.g. cancer, hypertension, atherosclerosis,
 CC cardiomyopathy or bronchial asthma. The products of the invention can
 CC be used for gene therapy or in a vaccine ABX97008-ABX97185 are cdna
 CC fragments amplified and isolated by the PCR primers and probes
 CC represented in ABX13460-ABX13462 and ABX97186-ABX97593. ABX97008-ABX97185
 CC encode the NOVX proteins described in ABU65041-ABU65218.
 XX
 SQ Sequence 1272 BP; 182 A; 521 C; 356 G; 213 T; 0 other;
 Query Match 4.8%; Score 47.2; DB 24; Length 1272;
 Best Local Similarity 48.8%; Pred. No. 0.084;
 Matches 127; Conservative 0; Mismatches 133; Indels 0; Gaps 0;
 QY 256 CCAAGAACATCAAGTGGCCAGAGAGTGGTGAAGACCCAGCGCGCTCTTCAACC 315
 Db 711 CCCCAGACAGAGTGTATGTGGTGGCGGAACCCCGTGACCCGGGCGCATCTCGACT 770
 QY 316 TCAAGCTCAGCTCAAGTCCAGTCCAGACCTGCTCAGCTTCTGCGGGGCTCTCCA 375
 Db 771 AGGCCCCAGACGCTCTCCAGACCCCGGGCTCCAGCTTCCGCGCTTGGCTTCGGCC 830
 QY 376 CTTAGTGTCCCATGTGGACAGTCCAGGCTACAGATGACTGGAGCTGTGGTCCAGC 435
 Db 831 ACGGCTGTGGGCCCCGTGGACACAGCTGGAGCGCGTCCGATCGGCTGTACGCCGAGC 890
 QY 436 CAGTGTCTGAGCTGGGCGCACTTCACTTCAGACAGAGGGCGAGGCCCGGAGTGG 495
 Db 891 ACCTTGACACTGGCTGGCGTACTTCCCGCTTCCCACTTCTGTCTGTCAGCGGGAGC 950
 QY 496 AGATGATCTGCCAGGCGTCC 515

Db 951 GTCTGGTCAGCGACCCGCGCC 970
 RESULT 24
 ABQ76896
 ID ABQ76896 standard; DNA; 53226 BP.
 XX
 AC ABQ76896;
 XX
 DT 13-MAR-2003 (first entry)
 XX
 DE Human G-protein coupled receptor DNA SEQ ID 3.
 KW G-protein coupled receptor; secretin receptor subfamily; human; SNP;
 KW GPCR; protease; Parkinson's disease; gene; chromosome X;
 KW single nucleotide polymorphism; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 CDS 3000..50651
 FT /*tag= a
 FT /product= "GPCR"
 FT /note "this coding sequence is interrupted by
 FT 13 introns"
 FT variation replace (1746,c)
 FT /*tag= b
 FT /note= "SNP, single nucleotide polymorphism"
 FT replace (1755,g)
 FT /*tag= c
 FT /note= "SNP, single nucleotide polymorphism"
 FT replace (1961,g)
 FT /*tag= d
 FT /note= "SNP, single nucleotide polymorphism"
 FT 3000..3088
 FT /*tag= e
 FT /number= 1
 FT intron 3089..3874
 FT /*tag= f
 FT /number= 1
 FT exon 3875..4038
 FT /*tag= g
 FT /number= 2
 FT intron 4039..6037
 FT /*tag= h
 FT /number= 2
 FT variation replace (5411,g)
 FT /*tag= i
 FT /note= "SNP, single nucleotide polymorphism"
 FT replace (5760,a)
 FT /*tag= j
 FT /note= "SNP, single nucleotide polymorphism"
 FT 6038..6170
 FT /*tag= k
 FT /number= 3
 FT intron 6171..8059
 FT /*tag= l
 FT /number= 3
 FT exon 8060..8178
 FT /*tag= m
 FT /number= 4
 FT intron 8179..15910
 FT /*tag= n
 FT /number= 4
 FT variation replace (11390,c)
 FT /*tag= o
 FT /note= "SNP, single nucleotide polymorphism"
 FT replace (11390,t)
 FT /*tag= p
 FT /note= "SNP, single nucleotide polymorphism"
 FT 15911..16127
 FT /*tag= q

FT	intron	/number= 5	
FT		16128..17484	
FT		/*tag= x	
FT	variation	/number= 5	
FT		replace (16988,a)	
FT		/*tag= s	
FT	exon	/note= "SNP, single nucleotide polymorphism"	
FT		17485..17647	
FT		/*tag= t	
FT	intron	/number= 6	
FT		17648..32332	
FT		/*tag= u	
FT	variation	/number= 6	
FT		replace (18361,g)	
FT		/*tag= v	
FT	variation	/note= "SNP, single nucleotide polymorphism"	
FT		replace (13769,c)	
FT		/*tag= w	
FT	variation	/note= "SNP, single nucleotide polymorphism"	
FT		replace (22910,t)	
FT		/*tag= x	
FT	variation	/note= "SNP, single nucleotide polymorphism"	
FT		replace (22935,t)	
FT		/*tag= y	
FT	variation	/note= "SNP, single nucleotide polymorphism"	
FT		replace (24206,g)	
FT		/*tag= z	
FT	variation	/note= "SNP, single nucleotide polymorphism"	
FT		replace (24774,a)	
FT		/*tag= aa	
FT	variation	/note= "SNP, single nucleotide polymorphism"	
FT		replace (24869,c)	
FT		/*tag= ab	
FT	variation	/note= "SNP, single nucleotide polymorphism"	
FT		replace (25766,a)	
FT		/*tag= ac	
FT	variation	/note= "SNP, single nucleotide polymorphism"	
FT		replace (26697,t)	
FT		/*tag= ad	
FT	variation	/note= "SNP, single nucleotide polymorphism"	
FT		replace (26697,c)	
FT		/*tag= ae	
FT	variation	/note= "SNP, single nucleotide polymorphism"	
FT		replace (28359,c)	
FT		/*tag= af	
FT	variation	/note= "SNP, single nucleotide polymorphism"	
FT		replace (28470..28471,t)	
FT		/*tag= ag	
FT		/note= "a single nucleotide polymorphism (SNP) can result in a deletion at this position"	
FT	variation	replace (29781,g)	
FT		/*tag= ah	
FT	variation	/note= "SNP, single nucleotide polymorphism"	
FT		replace (30182,a)	
FT		/*tag= ai	
FT	variation	/note= "SNP, single nucleotide polymorphism"	
FT		replace (31772,t)	
FT		/*tag= aj	
FT	variation	/note= "SNP, single nucleotide polymorphism"	
FT		replace (31936,g)	
FT		/*tag= ak	
FT	exon	/note= "SNP, single nucleotide polymorphism"	
FT		32333..32467	
FT		/*tag= al	
FT	intron	/number= 7	
FT		32468..36825	
FT		/*tag= am	
FT	exon	/number= 7	
FT		36826..36948	
FT		/*tag= an	
FT	intron	/product= 8	
FT		36949..38129	
FT		/*tag= ao	

FT		/number= 8	
FT	exon	38130..38175	
FT		/*tag= ap	
FT		/number= 9	
FT	intron	38176..39984	
FT		/*tag= aq	
FT		/number= 9	
FT	exon	39985..40088	
FT		/*tag= ar	
FT		/number= 10	
FT	intron	40089..42455	
FT		/*tag= as	
FT		/number= 10	
FT	exon	42456..42577	
FT		/*tag= at	
FT		/number= 11	
FT	intron	42578..44422	
FT		/*tag= au	
FT		/number= 11	
FT	variation	replace (42767..42767,c)	
FT		/*tag= av	
FT		/note= "a single nucleotide polymorphism (SNP) can result in a deletion at this position"	
FT	exon	44423..44691	
FT		/*tag= aw	
FT		/number= 12	
FT	intron	44692..47818	
FT		/*tag= ax	
FT		/number= 12	
FT	exon	47819..47897	
FT		/*tag= ay	
FT		/number= 13	
FT	intron	47898..50266	
FT		/*tag= az	
FT		/number= 13	
FT	variation	replace (48839,t)	
FT		/*tag= ba	
FT		/note= "SNP, single nucleotide polymorphism"	
FT	exon	50267..50651	
FT		/*tag= bb	
FT		/number= 14	
FT	variation	replace (52265,g)	
FT		/*tag= bc	
FT		/note= "SNP, single nucleotide polymorphism"	
XX	US2002142951-A1.		
PN			
PD	03-OCT-2002.		
XX			
FF	28-MAR-2001; 2001US-0818264.		
XX			
PR	28-MAR-2001; 2001US-0818264.		
XX			
PA	(WEBS/) WEBSTER M.		
PA	(BEAS/) BEASLEY E M.		
PA	(KETCH/) KETCHUM K A.		
PA	(DFRA/) DI FRANCESCO V.		
XX			
XX	Webster M, Beasley EM, Ketchum KA, Di Francesco V;		
PI			

RESULT 25
 AAX83003/c
 ID AAX83003 standard; DNA; 87350 BP.
 XX AC AAX83003;
 XX XX
 XX DT 31-AUG-1999 (first entry)
 XX XX
 XX DE Human WRN genomic sequence.
 XX XX
 XX KW Human; WRN; Werner's syndrome; detection; diagnosis; autosomal;
 XX KW recessive disorder; phenotype; ss.
 XX OS Homo sapiens.
 XX XX
 XX PN WO9724435-A1.
 XX PD 10-JUL-1997.
 XX XX
 XX PF 30-DEC-1996; 96WO-US20785.
 XX XX
 XX PR 12-APR-1996; 96US-0632175.
 XX PR 29-DEC-1995; 95US-0009409.
 XX PR 29-DEC-1995; 95US-0580539.
 XX PR 30-JAN-1996; 96US-0010835.
 XX PR 30-JAN-1996; 96US-0594242.
 XX XX
 XX PA (DARW-) DARWIN MOLECULAR CORP.
 XX PA (OSHI/) OSHIMA J.
 XX XX
 XX PI Fu Y, Mulligan J, Oshima J, Schellenberg GD, Yu C;
 XX XX
 XX DR WPI; 1997-363671/33.
 XX XX
 XX PT Isolated nucleic acid molecule encoding the WRN gene product -
 XX PT useful for detection and treatment of Werner's syndrome, and related
 XX PT diseases
 XX XX
 XX PS Claim 1; Fig 5A-U; 153pp; English.
 XX XX
 XX CC This sequence represents the genomic region containing the coding
 XX CC sequence for the human WRN gene which encodes a protein related to
 XX CC Werner's syndrome. The products can be used for the detection and
 XX CC treatment of Werner's syndrome (WS), an autosomal recessive disorder
 XX CC with a complex phenotype, as well as related diseases.
 XX XX
 XX SQ Sequence 87350 BP; 25621 A; 16221 C; 17012 G; 28450 T; 46 other;
 Query Match 4.7%; Score 47; DB 18; Length 87350;
 Best Local Similarity 78.9%; Pred. No. 0.39;
 Matches 56; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
 QY 922 GTATTGGAGTTCATGCAAAATGAGTGTGTTTGTAGTCTGTGCGCACAACAAAAA 981
 DB 18894 GTATTGGAGTTCCTGCTAAATGAGTGGATTTCAGCTGTTTGCCACAAAAA 18835
 QY 982 AAAAAA 992
 DB 18834 ATGGTTAACAA 18824

RESULT 26
 AAS68336/c
 ID AAS68336 standard; cDNA; 636 BP.
 XX AC AAS68336;
 XX XX
 XX DT 13-FEB-2002 (first entry)
 XX XX
 XX DE DNA encoding novel human diagnostic protein #4140.
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX Homo sapiens.
 XX PN WO200175067-A2.
 XX XX
 XX PD 11-OCT-2001.
 XX XX
 XX PF 30-MAR-2001; 2001WO-US08631.
 XX XX
 XX PR 31-MAR-2000; 2000US-0540217.
 XX PR 23-AUG-2000; 2000US-0649167.
 XX XX
 XX PA (HYSE-) HYSEQ INC.
 XX XX
 XX PI Drmanac RT, Liu C, Tang YT;
 XX XX
 XX DR WPI; 2001-639362/73.
 XX DR P-PSDB; ABG04149.
 XX XX
 XX PT New isolated polynucleotide and encoded polypeptides, useful in
 XX PT diagnostics, forensics, gene mapping, identification of mutations
 XX PT responsible for genetic disorders or other traits and to assess
 XX PT biodiversity -
 XX XX
 XX PS Claim 1; SEQ ID NO 4140; 103pp; English.
 XX XX
 XX CC The invention relates to isolated polynucleotide (I) and
 XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 XX CC and gene mapping, and in recombinant production of (II). The
 XX CC polynucleotides are also used in diagnostics as expressed sequence tags
 XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
 XX CC to restore normal activity of (II) or to treat disease states involving
 XX CC (II). (II) is useful for generating antibodies against it, detecting or
 XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
 XX CC imaging of sites expressing (II). (I) and its binding partners are useful in medical
 XX CC disorders involving aberrant protein expression or biological activity.
 XX CC The polypeptide and polynucleotide sequences have applications in
 XX CC diagnostics, forensics, gene mapping, identification of mutations
 XX CC responsible for genetic disorders or other traits to assess biodiversity
 XX CC and to produce other types of data and products dependent on DNA and
 XX CC amino acid sequences. AAS64197-AAS94564 represent novel human
 XX CC diagnostic coding sequences of the invention.
 XX CC Note: The sequence data for this patent did not appear in the printed
 XX CC specification, but was obtained in electronic format directly from WIPO
 XX CC at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 636 BP; 103 A; 246 C; 180 G; 107 T; 0 other;
 Query Match 4.5%; Score 44.4; DB 23; Length 636;
 Best Local Similarity 88.9%; Pred. No. 0.33;
 Matches 48; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 57 CGAGCATGGGGCTCCCTGGGCTGTTCTGTTGGCGTGTGCGTGCACGACC 110
 DB 263 CGAGCATGGGGCTCCCTGGGCTGTTCTGTTGGCGTGTGCGTGCACGACC 210

RESULT 27
 ABZ80229
 ID ABZ80229 standard; DNA; 249999 BP.
 XX AC ABZ80229;
 XX XX
 XX DT 02-JUN-2003 (first entry)
 XX XX
 XX DE Human transdormin gene region genomic DNA SEQ ID NO:26.
 XX KW Neuroprotective; nootropic; cerebroprotective; analgesic; gene therapy;
 XX KW central nervous system disorder; CNS disorder; multiple sclerosis;
 XX KW nerve injury; neuropathic pain; stroke; trauma; non-CNS disorder; tram;

```
KW tramodxin; human; chromosome 5; gene; ds.
XX
OS Homo sapiens.
XX
FN WO2003016502-A2.
XX
PD 27-FEB-2003.
XX
PF 21-AUG-2002; 2002WO-US26637.
XX
PR 21-AUG-2001; 2001US-313907P.
XX
PR 21-AUG-2002; 2002US-0225810.
XX
PA (MCLA-) MCLAUGHLIN RES INST.
XX
PI Birmingham JR;
XX
DR WPI; 2003-278567/27.
XX
XX New nucleic acid sequence encoding tramdorins, e.g. mouse tramd 1,
PT mouse tramd 2, mouse tramd 3, human tramd 1, human tramd 2, human tramd
PT 3 or rat tramd 1, useful for treating CNS, e.g. stroke, multiple
PT sclerosis, trauma, neuropathic pain -
XX
PS Example 6; Fig 9; 177pp; English.
XX
CC The present invention describes an isolated nucleic acid sequence
CC comprising a cDNA sequence encoding mouse tramdorin (tramd) 2, mouse
CC tramd 3, human tramd 1, human tramd 2, human tramd 3 or rat tramd 1, or
CC the genomic sequence of mouse tramd 1 or mouse tramd 3. Mouse tramd 1 is
CC located to chromosome 11, whereas human tramd 1 is located to chromosome
CC 5q31-33. The tramd sequences have neuroprotective, nootropic, analgesic
CC and cerebroprotective activities, and can be used in gene therapy. The
CC nucleic acid sequences are useful for diagnosing and treating central
CC nervous system (CNS) disorders such as multiple sclerosis, nerve injury,
CC neuropathic pain, stroke or trauma, and non-CNS disorders. The present
CC sequence represents the genomic sequence of the human tramdorin gene
CC region, which is given in the exemplification of the present invention.
XX
SQ Sequence 249999 BP; 75050 A; 54012 C; 51931 G; 67805 T; 1201 other;

Query Match 4.4%; Score 44; DB 25; Length 249999;
Best Local Similarity 77.9%; Pred. No. 3.1;
Matches 53; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 922 GTATTGGAGTTTCATGCAAAATGAGTGCTTTAGCTGCTCTTGCACAAAAA 981
Db 75447 GTATTAAAAATCTGTGAATGAGTAAATTTAGCTGCTCTTGCACAAAAA 75506

QY 982 AAAAAAAA 989
Db 75507 CCAATAA 75514

RESULT 28
ABT07673
ID ABT07673 standard; cDNA; 560 BP.
AC ABT07673;
XX ABT07673;
XX 14-NOV-2002 (first entry)
XX Human breast cancer associated coding sequence SEQ ID NO: 152.
DE Human; breast specific gene; breast specific protein; breast cancer;
KW gene therapy; cytostatic; gene; ss.
XX Homo sapiens.
XX WO200264611-A1.
XX 22-AUG-2002.

Query Match 4.4%; Score 44; DB 25; Length 249999;
Best Local Similarity 77.9%; Pred. No. 3.1;
Matches 53; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 924 ATTTGGAGTTTCATGCAAAATGAGTGCTTTAGCTGCTCTTGCACAAAAA 983
Db 408 ATATGGGATTCATCTAAATGAGTAGATTATTTAGCTGCTCTTGCACAAAAA 467

QY 984 AAAAAAAA 992
Db 468 ATGCCACAA 476

RESULT 29
ABT07672
ID ABT07672 standard; cDNA; 509 BP.
AC ABT07672;
XX ABT07672;
XX 14-NOV-2002 (first entry)
XX Human breast cancer associated coding sequence SEQ ID NO: 151.
DE Human; breast specific gene; breast specific protein; breast cancer;
KW gene therapy; cytostatic; gene; ss.
XX Homo sapiens.
XX WO200264611-A1.
XX 22-AUG-2002.

PF 12-FEB-2002; 2002WO-US04197.
XX
PR 13-FEB-2001; 2001US-268292P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Salceda S, Macina RA, Hu P, Recipon H, Karra K, Cafferkey R;
PI Sun Y, Liu C;
XX
DR WPI; 2002-657582/70.
XX
XX New breast specific nucleic acids and proteins, useful for identifying,
PT diagnosing, staging, imaging, and treating breast cancer
PT and non-cancerous disease states in breast tissue, and in gene therapy
PT -
XX
PS Claim 1; Page 269; 367pp; English.
XX
CC The present invention provides human breast specific coding sequences and
CC proteins. These can be used in the diagnosis and treatment of breast
CC cancer and non-cancerous diseases of the breast. The present sequence is
CC a coding sequence of the invention.
XX
SQ Sequence 560 BP; 173 A; 109 C; 139 G; 139 T; 0 other;

Query Match 4.4%; Score 43.4; DB 24; Length 560;
Best Local Similarity 76.8%; Pred. No. 0.56;
Matches 53; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 924 ATTTGGAGTTTCATGCAAAATGAGTGCTTTAGCTGCTCTTGCACAAAAA 983
Db 408 ATATGGGATTCATCTAAATGAGTAGATTATTTAGCTGCTCTTGCACAAAAA 467

QY 984 AAAAAAAA 992
Db 468 ATGCCACAA 476
```

tramodxin; human; chromosome 5; gene; ds.

Homo sapiens.

WO2003016502-A2.

27-FEB-2003.

21-AUG-2002; 2002WO-US26637.

21-AUG-2001; 2001US-313907P.

21-AUG-2002; 2002US-0225810.

(MCLA-) MCLAUGHLIN RES INST.

Birmingham JR;

WPI; 2003-278567/27.

New nucleic acid sequence encoding tramdorins, e.g. mouse tramd 1, mouse tramd 2, mouse tramd 3, human tramd 1, human tramd 2, human tramd 3 or rat tramd 1, useful for treating CNS, e.g. stroke, multiple sclerosis, trauma, neuropathic pain -

Example 6; Fig 9; 177pp; English.

The present invention describes an isolated nucleic acid sequence comprising a cDNA sequence encoding mouse tramdorin (tramd) 2, mouse tramd 3, human tramd 1, human tramd 2, human tramd 3 or rat tramd 1, or the genomic sequence of mouse tramd 1 or mouse tramd 3. Mouse tramd 1 is located to chromosome 11, whereas human tramd 1 is located to chromosome 5q31-33. The tramd sequences have neuroprotective, nootropic, analgesic and cerebroprotective activities, and can be used in gene therapy. The nucleic acid sequences are useful for diagnosing and treating central nervous system (CNS) disorders such as multiple sclerosis, nerve injury, neuropathic pain, stroke or trauma, and non-CNS disorders. The present sequence represents the genomic sequence of the human tramdorin gene region, which is given in the exemplification of the present invention.

Sequence 249999 BP; 75050 A; 54012 C; 51931 G; 67805 T; 1201 other;

Query Match 4.4%; Score 44; DB 25; Length 249999; Best Local Similarity 77.9%; Pred. No. 3.1; Matches 53; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

922 GTATTGGAGTTTCATGCAAAATGAGTGCTTTAGCTGCTCTTGCACAAAAA 981

75447 GTATTAAAAATCTGTGAATGAGTAAATTTAGCTGCTCTTGCACAAAAA 75506

982 AAAAAAAA 989

75507 CCAATAA 75514

RESULT 28

ABT07673

ID ABT07673 standard; cDNA; 560 BP.

ABT07673;

14-NOV-2002 (first entry)

Human breast cancer associated coding sequence SEQ ID NO: 152.

Human; breast specific gene; breast specific protein; breast cancer;

gene therapy; cytostatic; gene; ss.

Homo sapiens.

WO200264611-A1.

22-AUG-2002.

XX New human breast and ovarian cancer associated gene sequences and the
PT polypeptides encoded by these genes, useful in the prevention,
PT treatment and diagnosis of cancer, immune disorders, cardiovascular
PT disorders and neurological diseases -
XX
PS Claim 1; Page 555-556; 1299pp; English.
XX
XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
CC proteins AAB59111 - AAB59128. The DNA and protein sequences are
CC associated with breast and ovarian cancer. Included in the invention are
CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
CC isolation and characterisation of the DNA and protein sequences of the
CC invention. The breast and ovarian cancer associated DNA, protein, agonist
CC or antagonist sequences exhibit cytostatic; immunosuppressive;
CC neurotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
CC antidiabetic; antiinflammatory; antiulcer; vulnerary; anticonvulsant;
CC antibacterial; antifungal; antiparasitic and cardiant activity. The
CC polynucleotide and protein sequences are used in the diagnosis of cancer,
CC particularly breast and ovarian cancer. The nucleic acid sequences,
CC proteins, agonists and antagonists may also be used in the diagnosis,
CC prevention and treatment of immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; cardiovascular disorders such as
CC myocardial ischaemias; wound healing; neurological diseases such as
CC cerebral anoxia and epilepsy; and infectious diseases.
XX
SQ Sequence 752 BP; 149 A; 275 C; 219 G; 104 T; 5 other;
Query Match 4.2%; Score 41.8; DB 21; Length 752;
Best Local Similarity 51.3%; Pred. NO. 1.5;
Matches 97; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
QY 410 GATGCACTGGAGCTGTGTCTCAAGCCAGTCTCTGAGCTGGCGCCCACTTCACTCTGCA 469
Db 260 GCTGGCTGGCAGATGAGTCCACGGCGCTGTGTAGCGCAGAGACGTGGCCAGACC 201
QY 470 GGACAGAGGGCAGCGCCCGCCAGACGGCGCGCCCTGTCGCGCAGTGCAGCCACC 529
Db 200 CCGCAGAGGGCGCGCCCGCCAGACGGCGCGCCCTGTCGCGCAGTGCAGCCACC 141
QY 530 TATCACCACCAACCGCTTATGGCCATCGCGCTTGGGGGGCGAGTAGCGGTCTCTCACCGA 81
Db 140 TCCACACCAACCGCTTATGGCCATCGCGCTTGGGGGGCGAGTAGCGGTCTCTCACCGA 81
QY 590 CAGGCAGCC 598
Db 80 CAGGCAGGC 72
RESULT 32
AAC77293/c
ID AAC77293 standard; cDNA; 1417 BP.
XX
AC AAC77293;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF2848 polynucleotide sequence SEQ ID NO:5695.
XX
XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnerary; antiparasitic; antiparkinsonian; neurotropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;

KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive; ss.
XX
OS Homo sapiens.
XX
FN WO200058473-A2.
XX
PD 05-OCT-2000.
XX
PD 31-MAR-2000; 2000WO-US08621.
XX
PF 31-MAR-1999; 99US-0127607.
XX
PR 02-APR-1999; 99US-0127636.
XX
PR 05-APR-1999; 99US-0127728.
XX
PR 30-MAR-2000; 2000US-0540763.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
PI WPI; 2000-602362/57.
XX
DR P-PSDB; AAB43084.
XX
DR Novel nucleic acids and peptides derived from open reading frame X,
XX useful for treating e.g. cancers, proliferative disorders,
XX neurodegenerative disorders and cardiovascular disease -
PS Claim 5; Page 4858-4859; 5507pp; English.
XX
CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC antiparasitic; antiparkinsonian; neurotropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antithyroid; and antianaemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ Sequence 1417 BP; 211 A; 499 C; 473 G; 234 T; 0 other;
Query Match 4.2%; Score 41.8; DB 21; Length 1417;
Best Local Similarity 51.3%; Pred. NO. 1.9;
Matches 97; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
QY 410 GATGCACTGGAGCTGTGTCTCAAGCCAGTCTCTGAGCTGGCGCCCACTTCACTCTGCA 469
Db 987 GCTGGCTGGCAGATGAGTCCACGGCGCTGTGTAGCGCAGAGACGTGGCCAGACC 928
QY 470 GGACAGAGGGCAGCGCCCGCCAGACGGCGCGCCCTGTCGCGCAGTGCAGCCACC 529
Db 927 CCGCAGAGGGCGCGCCCGCCAGACGGCGCGCCCTGTCGCGCAGTGCAGCCACC 868
QY 530 TATCACCACCAACCGCTTATGGCCATCGCGCTTGGGGGGCGAGTAGCGGTCTCTCACCGA 599
Db 867 TCCACACCAACCGCTTATGGCCATCGCGCTTGGGGGGCGAGTAGCGGTCTCTCACCGA 808
QY 590 CAGGCAGCC 598
Db 807 CAGGCAGGC 799

RESULT 33
AAC77637
ID AAC77637 standard; cDNA; 2059 BP.
XX
AC AAC77637;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human cancer associated gene sequence SEQ ID NO:31.
XX
KW Human; cancer associated gene; cancer antigen; detection; cancer;
diagnosis; cytostatic; proliferative; vulnary; immunomodulator;
XX
KW antidiabetic; antiaesthetic; antirheumatic; antiarthritic; antiviral;
XX
KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
XX
KW dermatological; neuroprotective; thrombolytic; coagulant; neotropic;
XX
KW vasotropic; antiporiatic; antiangiogenic; gene therapy; inflammation;
XX
KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
XX
KW allergic reaction; graft versus host disease; organ rejection;
XX
KW haemostatic; thrombolytic; cardiovascular disorder; infection;
XX
KW neurological disease; drug screening; ss.
OS Homo sapiens.
XX
FN WO200055350-A1.
XX
PD 21-SEP-2000.
XX
PP 08-MAR-2000; 2000WO-US05982.
XX
PR 12-MAR-1999; 99US-0124270.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2000-587533/55.
XX
DR P-PSDB; AAB43428.
XX
PT Novel isolated nucleic acids comprising sequences encoding peptides
PT useful for treating or diagnosing e.g. cancer -
XX
PS Claim 1; Page 633-634; 2352pp; English.
XX
CC AAC77607 to AAC78448 encode the human cancer associated proteins given
CC in AAB43398 to AAB44239. The proteins can have activities based on the
CC tissues and cells the genes are expressed in. Example of activities
CC include: cytostatic; proliferative; vulnary; immunomodulator;
CC antidiabetic; antiaesthetic; antirheumatic; antiarthritic;
CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC neotropic; vasotropic; antiporiatic and angiogenic. The
CC polynucleotides and polypeptides can be used for preventing, treating or
CC ameliorating medical conditions and diagnosing pathological conditions.
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC the present invention may be used to treat immune disorders by activating
CC or inhibiting the proliferation, differentiation or mobilisation of
CC immune cells; to treat disorders of haematopoietic cells, autoimmune
CC disorders, allergic reactions, graft versus host disease and organ
CC rejection, modulate haemostatic or thrombolytic activity, modulate
CC inflammation, cancers, cardiovascular disorders, neurological disease and
CC bacterial or viral infections. The peptides, nucleotides, antibodies,
CC agonists and antagonists may be also be used in drug screens. AAC78449 to
CC AAC78457 and AAB44240 represent sequences used in the exemplification of
CC the present invention.
XX
SQ Sequence 2059 BP; 505 A; 548 C; 552 G; 451 T; 3 other;
XX
Query Match 4.2%; Score 41.8; DB 21; Length 2059;
Best Local Similarity 61.5%; Pred. No. 2.2;
Matches 67; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
QY 884 CAGGCCAGAGGACTGAGGCCATCAGCGTGCATGTCGTATTGGAGTTCATGCACAAAT 943

Db 1945 CAGGCCAGAGTCTGCGAGTCTCTGCGATGTTCTGCTTGTGCCCCCTTGTGACCCACAA 2004
QY 944 GAGTGTGTTTAGCTGCTCTTCCACAAAAAAGAAAAAAGAAAAAAGAAAAA 992
Db 2005 TAAAGGCTGAGCTCTTATCTTCCAAAAAAGAAAAAAGAAAAAAGAAAAA 2053

RESULT 34
ABX63699/c
ID ABX63699 standard; cDNA; 2672 BP.
XX
AC ABX63699;
XX
DT 26-FEB-2003 (first entry)
XX
DE Human cDNA #699 differentially expressed in activated vascular tissue.
XX
KW Human; gene; ss; vascular tissue; cytostatic; atherosclerosis;
XX
KW cardiant; hypotensive; antidiabetic; gynaecological; vasotropic;
XX
KW cerebroprotective; gene therapy; vascular disease; cancer; coronary;
XX
KW artery disease; hypertension; diabetes; pre-eclampsia; restenosis;
XX
KW ischaemia-reperfusion injury; stroke;
XX
OS Homo sapiens.
XX
FN US2002137081-A1.
XX
PD 26-SEP-2002.
XX
PP 08-JAN-2002; 2002US-0044090.
XX
PR 28-JUL-2000; 2000US-222469P.
XX
PR 08-JAN-2001; 2001US-260483P.
XX
PA (BAND/) BANDMAN O.
XX
PI Bandman O;
XX
DR WPI; 2003-110597/10.
XX
PT Combination for diagnosing, staging, treating, or monitoring the
PT progression of treatment of a vascular disease, e.g. atherosclerosis,
PT comprises several cDNAs that are differentially expressed in activated
PT vascular tissue -
XX
PS Claim 1; Page -; 18pp; English.
XX
CC This invention relates to a combination comprising several cDNAs that
CC are differentially expressed in activated vascular tissue. The invention
CC also discloses a high throughput method for detecting differentially
CC expressed cDNAs in a sample. The cDNAs of the invention may have
CC antiatherosclerotic; cytostatic; cardiant; hypotensive; antidiabetic;
CC gynaecological; vasotropic and cerebroprotective activities and may be
CC used in gene therapy. The cDNAs of the invention may be used in a
CC high-throughput methods for detecting differential expression of one or
CC more cDNAs in a sample, or screening several molecules or compounds to
CC identify a molecule or compound that specifically binds a cDNA of the
CC invention. A protein encoded by the cDNA may be used to screen several
CC molecules or compounds to identify a ligand that specifically binds to
CC the protein, or to produce or purify an antibody to the protein that can
CC be used to detect a protein in a sample or purify a natural or
CC recombinant protein from a sample. The nucleotides may be useful for
CC diagnosing, staging, treating, or monitoring the progression of
CC treatment of a vascular disease, e.g. atherosclerosis, cancer, coronary
CC artery disease, hypertension, diabetes, pre-eclampsia, ischaemia-
CC reperfusion injury, restenosis, or stroke. The cDNAs can also be used
CC for large-scale genetic or gene expression analysis of several new
CC nucleic acid molecules. Antibodies to the proteins encoded by the
CC cDNAs are useful for diagnosing pre-pathologic disorders, and chronic
CC or acute diseases associated with abnormalities in the expression,
CC amount or distribution of the protein. The present sequence
CC represents a cDNA of the invention that is differentially expressed in

RESULT 38
ABN97975
ID ABN97975 standard; DNA; 44100 BP.
XX
AC ABN97975;
XX
DT 01-AUG-2002 (first entry)
XX
DE Human retroviral sequence RAM75.
XX
KW Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
KW Multiple sclerosis; ds.
XX
OS Human retrovirus.
XX
PN WO9967395-A1.
XX
PD 29-DEC-1999.
XX
PF 23-JUN-1999; 99WO-FR01513.
XX
PR 23-JUN-1998; 98FR-0007920.
XX
PA (INERM) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Alliel PM, Perin J, Rieger F;
XX
DR WPI; 2000-160587/14.
XX
PT New nucleic acid sequences of human endogenous retrovirus, HERV-7q,
PT used for diagnosis, treatment and prevention of autoimmune and
PT neurological diseases -
XX
PS Claim 15; Fig 14; 225pp; French.
XX
CC The present invention relates to new nucleic acid sequences of human
CC endogenous retrovirus, HERV-7q, which is located on chromosome 7q.
CC Regulatory elements associated with HERV-7q may alter expression of other
CC genes (even remote genes) on the same chromosome, inducing immunological
CC and/or neurological changes (which may be pathological or protective/
CC curative). HERV-7q peptides can be used to improve efficiency of the
CC immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
CC sequences can be used in immunogenic or vaccinating compositions, for
CC protection against autoimmune diseases, particularly multiple sclerosis.
CC The peptides may also be used (by sequence comparison) to detect/identify
CC endogenous retroviruses that are abnormally expressed in cancer,
CC neuropathologies or other autoimmune diseases. The present sequence was
CC used to illustrate the invention.
XX
SQ Sequence 44100 BP; 13051 A; 8074 C; 8534 G; 14441 T; 0 other;
Query Match 4.0%; Score 39.6; DB 21; Length 44100;
Best Local Similarity 77.4%; Pred. No. 21;
Matches 48; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 920 TCGTATTGGAGTTCATCAAAATGAGTGTGTTTGGCTTGGCACAACAAAAA 979
DB 10427 TTGTATTGGGATTCATGTTAAATGAGTAGATTTTAACTACTCTTACCACAACACAA 10486
QY 980 AA 981
DB 10487 AA 10488
RESULT 39
ABN22420
ID ABN22420 standard; cDNA; 276 BP.
XX
AC ABN22420;
XX
DT 24-JUN-2002 (first entry)

XX
DE Human ORFX polynucleotide sequence SEQ ID NO:13317.
XX
KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KW hypertension; hypothyroidism; cholesterol ester storage disease;
KW immune deficiency; immune disorder; infectious disease;
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KW myasthenia gravis; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200192523-A2.
XX
PD 06-DEC-2001.
XX
PF 29-MAY-2001; 2001WO-US10836.
XX
PR 30-MAY-2000; 2000US-206132P.
XX
PR 29-AUG-2000; 2000US-228716P.
XX
PA (CURA-) CURAGEN CORP.
XX
XX Shimkets RA, Leach MD;
XX
XX WPI; 2002-106308/14.
XX
XX P-PSDB; ABP06668.
XX
PT Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and autoimmune disorders -
XX
PS Disclosure; SEQ ID 13317; 1037pp; English.
XX
CC The present invention describes substantially purified human proteins
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
CC in the specification). ABN15762 to ABN27252 encode the human ORFX
CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
CC treating or preventing a pathology associated with an ORFX-associated
CC disorder in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC sequences can be used in gene therapy. ORFX sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 276 BP; 43 A; 105 C; 86 G; 40 T; 2 other;
Query Match 4.0%; Score 39.4; DB 24; Length 276;
Best Local Similarity 50.8%; Pred. No. 4.4;
Matches 94; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
QY 338 TCCAGACCTGCTCACCTACTTCTGCGGGCGTCTTCCACCTCCAGTCCCATGTGGACAG 397
DB 9 TCCCCACCTGCTCGAGCGGCTCGACCTCGGCCACACACCCCTGGCAACCGGTGATCAT 68
QY 398 TGCAGGCTACAGATGACACTGGGAGCTGTGTGTCACGCCAGTGTCTGAGTGGGGCAA 457

Db 69 GGGCTCGATGCACACCGCATGGAGATCGTGCACAGCACCTGCCGAGCTGGCGGCTA 128
 QY 458 CTTCACTCTGCAGACAGAGGGGCGAGGCCCGCCAGGCTGGAGATGATCTGCCAGGCGTCTC 517
 Db 129 CTTCCGCGAGCGGGCCCAAGGCTGGCGCTGCTCGCTGCACCGCGCGCTACCGCCCAAC 188
 QY 518 GGGCA 522
 Db 189 TGGCA 193
 RESULT 40
 AAX80672
 ID AAX80672 standard; cDNA; 3285 BP.
 XX AC AAX80672;
 XX DT 22-OCT-1999 (first entry)
 XX DE Clone nm103_4 encoding secreted protein-nm103_4.
 XX KW Secreted protein; cytokine; cell proliferation; immune stimulation;
 KW vaccine; immune suppression; haematopoiesis; tissue growth; activin;
 KW inhibin; chemotaxis; chemokinesis; haemostasis; thrombolytic; ss;
 KW receptor; ligand; anti-inflammatory; cadherin; tumour; gene therapy.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 FT CDS 2294..2848
 FT /*tag= a
 FT /product= "Secreted protein nm103_4"
 XX PN W09928335-A1.
 XX XN 10-JUN-1999.
 XX PF 02-DEC-1998; 98WO-US25512.
 XX PR 30-NOV-1998; 98US-0203106.
 XX PR 04-DEC-1997; 97US-0067454.
 XX PA (GEMY) GENETICS INST INC.
 XX PI Agostino MJ, Clark HF, Collins-Racie LA, Evans C;
 PI Fechtel K, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
 PI Steininger RJ, Treacy M, Wong GG;
 XX WPI; 1999-385352/32.
 XX P-PSDB; AAV26037.
 XX NT New polynucleotides encoding secreted human proteins
 XX PS Claim 28; Pages 111-112; 124pp; English.
 XX CC The present sequence is a known clone nm103_4 (deposited as ATCC
 CC 98606) isolated from human foetal kidney cDNA library. It encodes
 CC secreted protein nm103_4. Recombinant secreted proteins can be produced
 CC by transfecting host cells and culturing them under suitable conditions.
 CC The polynucleotide and protein are predicted to have biological
 CC activities which would make them suitable for treating, preventing or
 CC ameliorating medical conditions in humans and animals. Some predicted
 CC biological activities include cytokine and cell proliferation/
 CC differentiation activity, immune stimulating (e.g. as vaccines) or
 CC suppressing activity, haematopoiesis regulating activity, tissue growth
 CC activity, activin/inhibin activity, chemotactic/chemokinetic activity,
 CC hemostatic and thrombolytic activity, receptor/ligand activity, anti-
 CC inflammatory activity, cadherin/tumour invasion suppressor activity, and
 CC tumour inhibition activity. The polynucleotide can be used for gene
 CC therapy.
 XX SQ Sequence 3285 BP; 756 A; 813 C; 705 G; 1011 T; 0 other;

Query Match 4.0%; Score 39.4; DB 20; Length 3285;
 Best Local Similarity 72.5%; Pred. No. 10;
 Matches 66; Conservative 0; Mismatches 21; Indels 4; Gaps 1;
 QY 891 GAGGACTGCGAGGCATCAGCGTGCACCTGTTGCGATTTGGAGTTTCATGCAAAATGAGTGTG 950
 Db 3049 GAGGAATGTAGGTAATAAATTCACCTGTT-----TTGGGAGTTCTCTGCTAAATGACTAGA 3104
 QY 951 TTTTAGCTGCTCTTGGCCACCAAAAAA 981
 Db 3105 CTTCACTGCTCTTGGCCACCAAAATCTTAAA 3135
 RESULT 41
 AAS59222
 ID AAS59222 standard; cDNA; 3285 BP.
 XX AC AAS59222;
 XX DT 16-JAN-2002 (first entry)
 XX DE Human cDNA encoding a secreted protein nm103_4.
 XX KW Human; secreted protein; ss; antiinflammatory; immunosuppressive;
 KW neotropic; neuroprotective; antiarthritic; antimicrobial; vulnerary;
 KW cyostatic; antidiabetic; virucide; antiinfertility; anticonvulsant;
 KW vasotropic; antiparkinsonian; immunostimulant; dermatological;
 KW antirheumatic; antitumor; antiulcer; osteopathic; tranquiliser;
 KW cerebroprotective; cytokine; cell proliferation; cell differentiation;
 KW immune deficiency; severe combined immunodeficiency; SCID; tumour;
 KW autoimmune disorder; multiple sclerosis; rheumatoid arthritis;
 KW graft-versus-host disease; myeloid deficiency; wound healing; ulcer;
 KW periodontal disease; osteoporosis; osteoarthritis; Alzheimer's disease;
 KW Parkinson's disease; Huntington's disease; infection; cardiac disease;
 KW stroke; sepsis; inflammatory bowel disease; contraceptive; immunogen;
 XX food supplement.
 XX OS Homo sapiens.
 XX WO200175068-A2.
 XX 11-OCT-2001.
 XX 22-MAR-2001; 2001WO-US09369.
 XX 30-MAR-2000; 2000US-0539330.
 XX 04-DEC-2000; 2000US-0729674.
 XX (GEMY) GENETICS INST INC.
 XX Jacobs K, McCoy JM, Lavallie E, Collins-racie LA, Evans C;
 XX Treacy M, Agostino MJ, Steininger RJ, Spaulding V, Wong GG;
 XX Clark H, Fechtel K, Merberg D;
 XX WPI; 2001-639363/73.
 XX P-PSDB; AAU39004.
 XX Secreted human proteins, useful as vaccine for treating various
 XX diseases such as autoimmune disorders (e.g. multiple sclerosis), and
 XX nervous system disorders (e.g. stroke) -
 XX Disclosure; Page 476-477; 619pp; English.
 XX The invention relates to novel human secreted proteins, the nucleic
 XX acids encoding them. The protein may exhibit cytokine, cell proliferation
 XX or cell differentiation activity or may induce production of other
 XX cytokines in certain cell populations and may exhibit immune stimulating
 XX or immune suppressing activity, which is useful for the treatment of
 XX various immune deficiencies and disorders e.g. severe combined
 XX immunodeficiency (SCID), autoimmune disorders e.g. multiple sclerosis,
 XX systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary
 XX inflammation. The proteins are also useful in the treatment of diseases

CC	with familial hypercholesterolaemia, hypoxanthine-guanine
CC	phosphoribosyltransferase associated with Lesch-Nyhan syndrome,
CC	phenylalanine hydroxylase (PAH) gene associated with phenylketonuria,
CC	dystrophin gene associated with muscular dystrophy, and human cystic
CC	fibrosis transmembrane conductance regulator gene associated with cystic
CC	fibrosis. The present nucleic acid sequence represents a human disease
CC	gene sequence that was used in the methods of the invention.
XX	
SQ	Sequence 56737 BP; 15689 A; 11281 C; 11599 G; 18168 T; 0 other;
	Query Match 3.9%; Score 38.6; DB 24; Length 56737;
	Best Local Similarity 59.9%; Pred. No. 41;
	Matches 100; Conservative 0; Mismatches 64; Indels 3; Gaps 2;
QY	819 GGGGATGGGAGTCTCAGAGGACGCGAAAGCAGCAGCATGTGAATGAACCGTCGACAGAG 878
Db	22684 GGGGAGATGTAGCTCAGAGGATACAAAGTAGCAGATGTAGGATCAGACAGCTAGAAA 22625
QY	879 CCAAGCAGCGCAGAGCATCTCAGGCGCATCAGCGTCACCTGTTTCGTATTGGAGTTCAFGC 938
Db	22624 TATAATGTTCAACATCAAGTATATAGGTAATAAAATTTGTGCT-GTATTGGGATTCACAC 22566
QY	939 ABAATCAGTGTGTTTGTAGCTGCTCTTGCCACAAAAAIAAAAAAAAA 985
Db	22565 TAAATGA--GATTTTAAAGTCTCTTTGCCACCAAAACAAAAAIAAAAA 22521

RESULT 47	
ABQ77405	
ID	ABQ77405 standard; DNA; 185695 BP.
XX	
XX	
AC	ABQ77405;
XX	
XX	
DT	10-MAY-2003 (first entry)
XX	
XX	
DE	Human THBS1 DNA.
XX	
XX	
KW	Human; THBS1; vascular disease; cardiac; antiarteriosclerotic; stroke;
KW	cardioprotective; gene therapy; coronary artery disease; ischaemia;
KW	myocardial infarction; peripheral vascular disease; pulmonary embolism;
KW	venous thromboembolism; forensic; paternity testing; G112583762; gene;
KW	SNP; single nucleotide polymorphism; ds.
XX	
XX	
OS	Homo sapiens.

Key	Location/Qualifiers
variation	replace (53502,t)
	/*tag= a
	/standard_name= "SNP"
	/note= "Single nucleotide polymorphism (ID G334u3) which does not change the THBS1 protein"

XX	WO2003016494-A2.
XX	
XX	27- FEB-2003.
XX	
XX	16-AUG-2002; 2002WO-US26343.
XX	
XX	16-AUG-2001; 2001US-313097P.
PR	05-OCT-2001; 2001US-327485P.
PR	14-DEC-2001; 2001US-0020141.
XX	
XX	(VITI-) VITIVITY INC.
XX	
XX	Mccarthy J, Ableson A;
XX	
XX	WPI: 2003-300617/29.
DR	P-PSDB: AEG74673.
DR	
XX	
XX	Identifying a subject as a candidate for a particular course of therapy
PT	to treat a vascular disease or disorder, e.g. stroke, myocardial
PT	infarction or ischemia by determining the identity of the nucleotide
PT	present at specific positions -
PT	

XX	Claim 3; Fig 11; 568pp; English.
XX	This invention describes a novel method for identifying a subject as a
CC	candidate for a particular course of therapy to treat a vascular disease
CC	or disorder. The method comprises determining the identity of the
CC	nucleotide present at specific positions, or their complements, and
CC	identifying the subject as a candidate for a particular clinical course
CC	of therapy based on the identity of the nucleotide present in that
CC	specific position. The method can be used for identifying a subject who
CC	is a candidate for further diagnostic evaluation of a vascular disease or
CC	disorder and selecting a clinical course of therapy. The products of the
CC	invention have cardiant, antiarteriosclerotic and cerebroprotective
CC	activity and can be used for gene therapy. The methods disclosed are
CC	useful for treating a vascular disease, e.g. atherosclerosis, coronary
CC	artery disease, myocardial infarction, ischaemia, stroke, peripheral
CC	vascular diseases, venous thromboembolism and pulmonary embolism. The DNA
CC	sequences are useful as fingerprint for detecting different individuals
CC	within the same species applicable in forensic studies and paternity
CC	testing. This sequence encodes the human THBS1 gene represented in
CC	G112593762, used to illustrate the method of the invention.
XX	
SQ	Sequence 185695 BP; 59388 A; 38636 C; 35664 G; 52007 T; 0 other;
	Query Match 3.9%; Score 38.4; DB 25; Length 185695;
	Best Local Similarity 53.7%; Pred. No. 68;
	Matches 101; Conservative 0; Mismatches 86; Indels 1; Gaps 1;
QY	795 GAGAGATTGGGGGTTTCAGATAGGGAATGGGAGGTCTCAGAGACCGAAGCACGACC 854
Db	112437 GATGTGCTGAGACTGCTGAGAAATGGGGAGATATACATAAAGGATACAAGTTGCAGAC 112496
QY	855 ATGTPAGAATCAACCGTCCAGAGACCCAGACGCCGAGAGGACTGCAGGCCCATCAGCGTGCC 914
Db	112497 ATGTAAGATGAACAAAGTCTGAGGTCTAACATACAACTAGGACTATAGTTAGTAATAC 112556
QY	915 ACTGTTTCGATTTGGAGTTTCATGCAAAATGAGTGTGTTTTAGTCTCTCTGCCACAAAA 974
Db	112557 TGCCTT-TTATTTCGAGATCTTCGAAAAAGTAGATTTTACGTGTTCTTTCCACACATA 112615
QY	975 AAAAAAAAA 982
Db	112616 CGCAAAGA 112623
RESULT 48	
ABV61090/c	
ID	ABV61090 standard; cDNA; 294 BP.
XX	ABV61090;
XX	13-SEP-2002 (first entry)
XX	Human prostate expression marker cdna 61081.
DE	Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XW	pharmacogenomic marker; gene; ss.
KX	

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
DR
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer -
XX
XX Claim 1; Page 11605-11606; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for:
XX (a) assessing whether a patient is afflicted with prostate cancer;
XX (b) monitoring the progression of prostate cancer in a patient;
XX (c) assessing the efficacy of a test compound to inhibit prostate
XX cancer in a patient;
XX (d) assessing the efficacy of a therapy for inhibiting prostate cancer
XX in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound;
XX (g) determining whether prostate cancer has metastasized in a patient;
XX (h) assessing the aggressiveness or indolence of prostate cancer in a
XX patient;
XX (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
XX Sequence 294 BP; 104 A; 26 C; 69 G; 95 T; 0 other;
SQ

Query Match 3.8%; Score 38; DB 23; Length 294;
Best Local Similarity 81.5%; Pred. No. 9.9;
Matches 44; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 939 AAAATGAGTGTGTTTACGTCTTGTCCACAAAAAATAAAAAA 992
DB 93 AAAAGTTTGTGTTTACCCCTTCCCCAAAAAATAAAAAA 40
RESULT 49
ABQ81848
ID ABQ81848 standard; DNA; 349980 BP.
XX
XX ABQ81848;
AC
XX
XX 19-NOV-2002 (first entry)
DT
XX
XX Bifidobacterium longum NCC2705 related nucleotide sequence SEQ ID:1104.
DE
XX
XX Bifidobacterium longum NCC2705; Bifidobacterium; bacterial;
XX antidiarrheic; antibacterial; inhibitor of Salmonella; detection;
XX identification; lactic acid bacterium; diarrhoea; pathogenic bacteria;
XX rotavirus; food composition; pharmaceutical composition; gene; ds.
XX
XX Bifidobacterium longum.
OS
XX Synthetic.
XX
XX EP1227152-A1.
PN
XX
XX 31-JUL-2002.
PD
XX
XX 30-JAN-2001; 2001EP-0102050.
PF
XX
XX 30-JAN-2001; 2001EP-0102050.
PR
XX
XX (NEST) SOC PROD NESTLE SA.
PA
XX
XX WPI; 2002-668397/72.
DR
XX
XX Novel polynucleotide comprising Bifidobacterium genome sequence useful
XX as a probe or primer for detecting and/or identifying Bifidobacterium
XX longum in a biological sample -
XX

PS Disclosure; SEQ ID 1104; 80pp; English.
XX
XX The present invention describes a polynucleotide (I) comprising a
XX sequence of a Bifidobacterium genome selected from the nucleotide
XX sequences given in ABQ81842 and ABQ81843, or a sequence exhibiting at
XX least 90% identity or which hybridises with the sequences given in
XX ABQ81842 and ABQ81843. Also described is a polynucleotide (II) encoding
XX a fusion protein, comprising a sequence selected from 1097 sequences
XX given in ABP65258 to ABP66354 ligated in frame to a polynucleotide
XX encoding a heterologous polypeptide. (I) has antidiarrheic and
XX antibacterial activities, and can be used as an inhibitor of Salmonella.
XX (I) (which is a probe) is useful for the detection and/or identification
XX of Bifidobacterium longum in a biological sample. A carrier containing
XX the lactic acid bacterium Bifidobacterium longum NCC2705 (NCIM 1-2618)
XX can be used for preventing and/or treating diarrhoea brought about by
XX pathogenic bacteria and/or rotavirus. The carrier is a food composition
XX selected from milk, yogurt, curd, cheese, fermented milks, milk based
XX fermented products, ice-creams, fermented cereal based products, milk
XX based powders, infant formula, pet food or a pharmaceutical composition
XX selected from tablets, liquid bacterial suspensions, dried oral
XX supplement, wet oral supplement, dry tube feeding or wet tube feeding.
XX (I) is useful in DNA arrays or chips to carry out analysis of the
XX expression of the Bifidobacterium gene. ABQ81844 to ABQ81850 represent
XX Bifidobacterium related nucleotide sequences given in the sequence
XX listing from the present invention but not mentioned further within the
XX specification.
XX N.B. The sequence data for this patent is not represented in the printed
XX specification but is based on sequence information supplied by the
XX European Patent Office.
XX
XX Sequence 349980 BP; 69195 A; 106952 C; 106128 G; 67705 T; 0 other;
SQ

Query Match 3.8%; Score 38; DB 24; Length 349980;
Best Local Similarity 60.8%; Pred. No. 1e+02;
Matches 62; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY 255 ACCAAGAACATCAAGTGGCCAGAGAGTGTGAAGACCCACGAGCGGCTCTTCAAC 314
DB 222380 ACCAAGACCTTCAAGACCGACAGAGGTCCAGGCTTACTCCGAGCGCGTCCGTCATC 222439
QY 315 CTCACGCTCACACTCAAGTCCAGTCCAGCTGCTCACCTAC 356
DB 222440 CAGACCTTCAACATCCCGTCTCCCTTGACCACTTCAAGAAC 222481
RESULT 50
ABV18588
ID ABV18588 standard; cDNA; 353 BP.
XX
XX ABV18588;
AC
XX
XX 13-SEP-2002 (first entry)
DT
XX
XX Human prostate expression marker cDNA 18579.
DE
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX WO2001:60860-A2.
PN
XX
XX 23-AUG-2001.
PD
XX
XX 20-FEB-2001; 2001WO-US05171.
PF
XX
XX 17-FEB-2000; 2000US-183319P.
PR
XX
XX 16-MAR-2000; 2000US-189862P.
PR
XX
XX 25-MAY-2000; 2000US-207454P.
PR
XX
XX 09-JUN-2000; 2000US-211314P.
PR
XX
XX 18-JUL-2000; 2000US-219007P.
PR
XX
XX 13-DEC-2000; 2000US-255281P.
PR
XX

